USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ESCOLA DE DOUTORAMENTO
INTERNACIONAL

Michael
González
Durruthy

Tese de doutoramento

Theoretical-experimental study on protein-ligand interactions based on thermodynamics methods, molecular docking and perturbation models

Santiago de Compostela, 2021

PD en Ciencia de Materiais

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

TESE DE DOUTORAMENTO

THEORETICAL-EXPERIMENTAL STUDY ON PROTEIN-LIGAND INTRACTIONS BASED ON THERMODYNAMICS METHODS, MOLECULAR DOCKING AND PERTURBATION MODELS

Michael González Durruthy

ESCOLA DE DOUTORAMENTO INTERNACIONAL DA UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PROGRAMA DE DOUTORAMENTO EN CIENCIA DOS MATERIALES

SANTIAGO DE COMPOSTELA, 2021

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARACIÓN DEL AUTOR/A DE LA TESIS

D./Dña. **Michael González Durruthy**

Título da tese: Theoretical-experimental study on protein-ligand interactions based on thermodynamics methods, molecular docking and perturbation models

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento y declaro que:

1) La tesis abarca los resultados de la elaboración de mi trabajo.
2) De ser el caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.
3) Confirmo que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.

Y me comprometo a presentar el Compromiso Documental de Supervisión en el caso que el original no esté depositado en la Escuela.

En **Santiago de Compostela**, **03 de septiembre de 2021**.

Firma electrónica

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AUTORIZACIÓN DO DIRECTOR / TITOR DA TESE

Theoretical-experimental study on protein-ligand interactions based on thermodynamics methods, molecular docking and perturbation models.

D./Dna. Juan M. Ruso Veiras.

INFORMA/N:

Que a presente tese, correspóndese co traballo realizado por D/Dna. Michael González Durruthy , baixo a miña dirección e titorización, e autorizo a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como director desta non incorre nas causas de abstención establecidas na Lei 40/2015.

De acordo co indicado no Regulamento de Estudos de Doutoramento, declara tamén que a presente tese de doutoramento é idónea para ser defendida en base á modalidade de Monográfica con reproducción de publicaciones, nos que a participación do/a doutorando/a foi decisiva para a súa elaboración e as publicacións se axustan ao Plan de Investigación.

En Santiago de compostela, 14. de octubre. de 2021.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Agradecimientos.

Después de mucho andar, finalmente consigo escribir estas líneas con sentimiento similar al de un peregrino por los caminos de Santiago...

Acompañado de dosis renovables de motivación, ilusión y la convicción de que siempre se puede más, he podido vislumbrar el final del camino de esta ardua jornada de incertidumbres y desafíos…

Hoy deseo dedicar estas líneas de agradecimiento a cada una de las personas que me incentivaron en esta caminada, sin dudas hicieron mucho más leve mi equipaje en momentos decisivos para este peregrino académico por los caminos de Santiago...

Dedico esta tesis a mi padre y mi madre por sus valientes esfuerzos desde Cuba, por priorizar siempre mis estudios, por ser siempre mi fuente de inspiración…

A mi Director de tesis, absolutas gracias, por darme la oportunidad de luchar, por perseverar conmigo en la consecución de este sueño, por sus pertinentes y tranquilizadores consejos, y sobre todas las cosas por brindarme su amistad…

A mi esposa y compañera Diana, por ser incondicional a toda prueba, por ser testigo y cómplice de este sueño desde sus orígenes, por las horas de vigilia compartidas, por ser decisiva en el aspecto emocional para poder seguir adelante, esta conquista también es tuya Chiqui…

A mis amigas de la vieja guardia Anaely y Denisell, por brindarme todo su apoyo durante la tramitación de mis documentos en pleno auge de la pandemia en Cuba…

A Roberto Borrego por su amistad y ayuda incondicional, muchas gracias hermano…

A mi colega Ramón Rial, muchas gracias por tu disponibilidad y buenos deseos…

A mis abuelos Tito y Guillermina…

A la profesora Natália Cordeiro, por abrirme las puertas en el Departamento de Química y Bioquímica de la Faculdade de Ciências da Universidade do Porto…

A Humbert, por sus consejos y amistad sincera…

A la Escuela Internacional de Doctorado de la USC…

¡A todos y todas infinitas gracias!

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

“I am among those who think that science
has great beauty”
Marie Skłodowska Curie

Resumo

Comprender a termodinámica das interaccións proteína-ligante foi fundamental para o desenvolvemento e aplicacións actuais da Química Medicinal. Desde que en 1894 Hermann Emil Fischer definisse as bases do recoñecemento molecular a través da formulación do "principio da chave e a fechadura". O coñecemento achega do axuste xeométrico entre as partes interactuantes (proteína + ligante), constitúe un requisito previo o cal determina e exerce unha influencia directa na afinidade das interaccións. Con todo, a complementariedade xeométrica por si soa non contribúe á formación de complexos de axuste proteína-ligante estables. Neste contexto, as interaccións non covalentes constitúen a forza esencial que sustenta a afinidade de unión entre as moléculas interactuantes (proteína + ligante) para un recoñecemento molecular exitoso. Por suposto, unha aplicación rigorosa do "principio de chave e fechadura" de Emil Fisher deberá obrigatoriamente considerar a flexibilidade ou rixidez dos mencionados sistemas de axuste ou recoñecemento molecular (proteína + ligante). Ao mesmo tempo que deberá considerar criterios sobre o grao de complementariedade e balance entre as propiedades hidrofóbicas e hidrofílicas de proteínas e ligantes.

Estes estudos sentaron as bases para a comprensión dos mecanismos de interacción das encimas cos seus substratos específicos. Neste sentido, evidencias bioquímicas experimentais mostran que só unha pequena porción da encima, chamada sitio activo, participa na interacción cun ou varios substratos durante unha reacción bioquímica.

O sitio activo nunha encima é unha entidade tridimensional formada por grupos que proveñen de diferentes partes das cadeas polipeptídicas que conforman as proteínas únense mediante o encartado e a flexión das cadeas formando usualmente cavidades hidrofóbicas. Devandito sitio activo comprende unha rexión relativamente pequena da estrutura da encima a cal participa directamente na catálisis. Debido a que esta contén os requirimentos estruturais en termos de atributos estereoquímicos e conformacionales para garantir o axuste preciso dos substratos e/ou ligantes (fármacos) dentro as mencionadas cavidades. A combinación dunha encima e un substrato específico forma un complexo encima-substrato (ES), que proporciona un roteiro alternativo para a reacción cunha menor enerxía de activación. Dentro do complexo encima-substrato, o substrato atopa condicións de reacción máis favorables que se estivese libre, e o

resultado é unha formación máis rápida do produto.

Os recentes avances nas técnicas de cristalografía de raios X integrados con métodos de Bioloxía Estrutural e Computacional levaron a unha explosión significativa no número de estruturas publicamente dispoñibles en bancos de datos de proteínas (do inglés, "*Protein Data Bank*"), a cal recentemente superou as 123,456 estruturas proteicas cristalografiadas. "*Research Collaboratory Structural Bioinformatics Protein Data Bank* (RCSB PDB)", constitúe un dos maiores repositórios de datos de U.S.A para o arquivo global de estruturas modeladas PDB, servindo a decenas de miles de depositantes de datos ao redor do mundo. Os datos da estrutura macromolecular 3D para proteínas e acedos nucleicos están dispoñibles sen cargo e sen restricións para millóns de usuarios de RCSB.org en todo o mundo, incluídos 660,000+ educadores, estudantes e membros do público curioso que utilizan PDB101.RCSB.org. Os depositantes de datos de PDB inclúen biólogos estruturais que utilizan ferramentas avanzadas de cristalografía, espectroscopia de resonancia magnética nuclear, microscopia electrónica 3D e difracción de microelectrones para xerar modelos apropiados para o estudo estrutural e mecanístico de múltiples problemas de relevancia biolóxica. Os usuarios de datos PDB que acceden aos portais web inclúen investigadores, educadores e estudantes fundamentalmente de biomedicina, biotecnoloxía, bioingeniería e ciencias da enerxía. Durante os últimos 2 anos, o portal web RCSB PDB redeseñouse por completo, o que permite unha procura avanzada baseada en lóxica booleana co acceso máis fácil aos datos PDB integrados para máis de 40 recursos baseado en biodatos externos. En 2020 los archivos PDB de proteínas descargadas totalizaron ~840 millones (~ 2,3 millones de descargas de archivos/día)

Doutra banda, a base de datos denominada "*PubChem*" é unha das bases de datos máis recoñecidas e en continuo crecemento contendo máis de 15,5 millóns de entradas sobre compostos puros e mesturas, a cal permite a obtención de información cristalográfica relevante para o estudo teórico das propiedades de interacción dos compostos ligantes, (incluíndo os compostos derivados da nanotecnoloxía).

Doutra banda, a base de datos PubChem (https://pubchem.ncbi.nlm.nih.gov) é un repositorio público de información sobre ligantes (substancias químicas) e as súas actividades biolóxicas, lanzado en 2004 como iniciativa das Bibliotecas Moleculares e Institutos Nacionais de Saúde de E.U.A (NIH). Durante os últimos 11 anos, PubChem creceu significativamente ata converterse nun sistema de almacenamento de grandes volumes de datos relevantes sobre ligantes, que serve como recurso de información

química para toda a comunidade de científica. PubChem consta de tres basees de datos interconectadas, i) Substance, ii) Compound e iii) BioAssay. Substance e Compound contén información química estrutural relevante. Mentres os datos asociados á súa actividade biolóxica experimental atópanse na base de datos de BioAssay. Adicionalmente, PubChem ofrece unha breve descrición das estruturas teóricas tridimensionales de compostos, así como arquivos de modelos de estrutura 3D de ligantes nos formatos .SDF, JSON, .XML, ASNT que poden ser utilizados eficientemente en estudos téorico-preditivos. Nos últimos anos o contido da información espectral contida em PubChem aumentou considerablemente. Neste sentido, no presente 591,000+ compostos en almacenados en PubChem teñen un ou máis tipos de información espectral dispoñible co acceso aberto á comunidade científica internacional, incluíndo: ^{13}C NMR, ^{1}H NMR, 2D NMR, ATR-IR, FT-IR, MS, GC-MS, Raman, UV–Vis, IR en fase de vapor e moitos máis. Vale a pena notar que a información espectral de PubChem provén de SpectraBase (http://spectrabase.com), un repositorio espectral en liña gratuíto de Bio-Rad Laboratories, Inc. A cal proporcionou imaxes interactivas, anotacións e ligazóns a un amplo e diverso conxunto de información espectral para 225,000+ ligantes (incluíndo nanopartículas orgánicas como fulereno C60). Tamén ofrecendo datos sobre NMR, IR, RAMAN, UV e MS conxuntamente con metadatos sobre 173,000 espectros de masas para 77,000 compostos.

Neste contexto, os procesos termodinámicos constitúen a base fundamental dos eventos de recoñecemento molecular proteína-ligante os cales son esenciais para identificar potenciais aplicacións biomédicas (bioactividad de novas moléculas), xogando un papel central desde etapas moi temperás da investigación preclínica no desenvolvemento exitoso de novos principios activos (fármacos).

Paralelamente constatouse que o número e diversidade molecular de compostos (fármacos, nanopartículas, nanomateriales, etc.) experimentou actualmente un crecemento desproporcionado. Este feito é consecuencia directa do desequilibrio existente entre a velocidade á que se desenvolven e comercializánse os novos ligantes (fármacos) e os curtos tempos de avaliación rigorosa sobre os riscos farmacolóxicos (interaccións farmacolóxicas) e toxicolóxicos potenciais para a saúde humana e mesmo a seguridade ambiental.

O escenario vólvese aínda máis complexo se se considera que a maioría dos estudos mecanísticos de farmacoloxía e toxicoloxía experimental sobre interacción proteína-

ligante, adoitan ser moi custosos en termos de infraestrutura, recursos humanos e tempo. Igualmente, as consideracións bioéticas asociadas á investigación experimental con animais de laboratorio e o seu posterior sacrificio ao final de cada proxecto de investigación de novos fármacos gañaron cada vez maior interese.

Neste contexto, as institucións académicas, empresas farmacéuticas e a industria da biotecnoloxía desenvolveron novas e eficientes estratexias metodolóxicas para mellorar a avaliación das interaccións proteína-ligante orientadas ao deseño racional de fármacos en Química Medicinal. Devanditos estudos baséanse principalmente no uso de ferramentas da Química Computacional e Bioinformática aplicada ao estudo termodinámico de interaccións proteína-ligante (fármaco-receptor). Como exemplos representativos podemos citar as técnicas de: simulación de axuste molecular, cribado virtual, enfoques biofísico-mecanísticos utilizando modelos de perturbación local e modelos preditivos sobre relacións cuantitativas estrutura-actividade-toxicidade (modelos QSAR/QSTR). Sumándose recentemente á lista, a aplicación de modelos de aprendizaxe automática baseados en intelixencia artificial e redes neuronais arificiales.

Como complemento dos ensaios termodinámicos proteína-ligante, os métodos computacionales "*in silico*" han contribuído significativamente a proporcionar interpretacións mecanísticas sobre as interaccións proteína-ligante cunha precisión razoable, eficiencia preditiva e potencial para extrapolación de resultados teóricos a experimentais. Neste sentido, as simulacións teóricas utilizando modelos multiescala permiten o estudo das propiedades termodinámicas dos complexos proteína-ligante, o que tamén é moi útil para a predición de mecanismos de interacción a través dos distintos niveis de organización (atómico, molecular, celular).

Especificamente, os métodos termodinámicos de axuste molecular comunmente utilizan dous algoritmos por separado: i) un primeiro algoritmo de mostraxe e procura que #predicir todas as configuracións ou conformacións estruturais (poses) que o ligante pode adoptar dentro do sitio activo dunha proteína branco, e ii) un algoritmo baseado nunha función de puntuación que #predicir as enerxías de afinidade de unión entre o ligante e a proteína receptora para cada unha das configuracións establecidas polo algoritmo de procura e mostraxe. Logo, combinando estes dous algoritmos, todas as configuracións de unión do ligante e o seu receptor proteico son clasificadas (ou jerarquizadas termodinámicamente) de acordo ás súas enerxías (ou afinidades de unión). Por tanto, as funcións de puntuación son capaces de filtrar compostos ligantes a partir de grandes bases de datos nunha procura virtual, onde a configuración dun

ligante co máis alto puntaje debería corresponder á configuración cunha enerxía de unión máis favorable consecuente con procesos termodinámicos espontáneos, para que este ligante ou composto poida ser considerado como un potencial candidato cabeza de serie. A maioría dos programas de axuste molecular, son executados a través de diferentes métodos para tratar cada caso en particular. Debendo ser considerado que cada proteína receptora é diferente, e así como a súa capacidade de replicar os resultados experimentais e fisiolóxicos dependerá en gran medida do programa utilizado e da experiencia e criterios establecidos polo investigador.

Outra das aplicacións recente dos métodos de simulación de axuste molecular é a súa integración con técnicas de intelixencia artificial orientadas ao deseño racional de novos fármacos, as cales se utilizou con éxito para xerar novos candidatos experimentalmente sintetizables e con alta afinidade de unión a biomoléculas (proteínas) con potencial interese biomédico.

As simulacións axuste molecular, predín teoricamente a afinidade de interacción entre proteínas e ligantes (fármacos), e son amplamente aceptadas no deseño computacional de novos fármacos. As bases teóricas do proceso de recoñecemento entre un ligante e o seu receptor susténtanse na estrutura cristalográfica 3D e a conformación de mínima enerxía adoptada polo/os ligante(s) durante a interacción co receptor diana, seguindo os principios da teoría de "axuste inducido".

O modelo de “chave e fechadura” explica a acción de numerosas encimas. Con todo, é demasiado restritivo para a acción de moitas outras encimas. A evidencia experimental indica que moitas encimas teñen flexibilidade nas súas formas, non son ríxidas e estáticas, senón que hai un cambio constante na súa forma. O modelo de axuste inducido utilízase para este tipo de situación. O modelo de axuste inducido permite pequenos cambios na forma ou xeometría do sitio activo dunha encima para acomodar un substrato. Unha boa analoxía son os cambios que ocorren na forma dunha luva cando se insere unha man nel. O axuste inducido é o resultado da flexibilidade da encima, xa que se adapta para aceptar o substrato entrante. Este modelo é unha explicación máis exhaustiva das propiedades do sitio activo dunha encima porque inclúe a especificidade do modelo de “chave e fechadura” xunto coa flexibilidade da encima

Neste contexto, as técnicas de simulación de axuste molecular poden clasificarse de acordo a tres categorías fundamentais baseadas en criterios de flexibilidade para o receptor e/ou ligante: i) axuste ríxido, ii) axuste semiflexible e iii) axuste flexible. No

axuste ríxido, a estrutura dos elementos que interactúan (proteína e ligante) non varía en termos de distancias e ángulos de ligazóns, permanecendo ambos os parámetros constantes. O método é relativamente simple e avalía principalmente o grao de coincidencia conformacional, o que o fai máis adecuado para o estudo de sistemas macromoleculares (proteínas con múltiples domínos, ADN, outros). No segundo caso o axuste semiflexible, permite variar as propiedades de flexibilidade conformacional das dúas moléculas que interactúan (proteína e/ou ligante) dentro dun certo rango de graos de liberdade conformacional, o cal o fai máis apropiado para o estudo de interaccións entre proteínas e ligantes pequenos. Para este caso a estrutura un determinado ligante pódese cambiar libremente mentres que a estrutura do receptor permanece ríxida ou nalgúns casos algúns dos residuos de aminoácidos do receptor (proteína) son definidos como rotables para optimizar o custo computacional durante o axuste molecular. Na terceira categoría, o axuste flexible permite o estudo íntegro das propiedades de flexibilidade conformacional do sistema (é dicir, proteína e ligando simultaneamente), os cales se poden avaliar libremente considerando todos os graos de liberdade rotacional e conformacional. Este último tipo de axuste molecular consome máis recursos computacionales, con todo, mellora significativamente a precisión e a sensibilidade das simulacións.

Os capítulos que compoñen a presente tese doutoral estruturáronse en nove secciones centrais cos seus sub-seccións correspondentes, as cales describen aspectos clave sobre a termodinámica e interaccións inter-moleculares proteína-ligante. Abarcando desde o enfoque tradicional da Química Medicinal ata o campo emerxente e en contínuo crecemento da Ciencia de Materiais (Nanotecnoloxía). Especificamente, no contexto das aplicacións biomédicas das nanopartículas e nanomateriales de carbono.

Actualmente, evidénciase un avance significativo na comprensión das forzas termodinámicas implicadas nas interaccións proteína- ligante. Aquí serán presentados varios exemplos que abordan a complexidade das interaccións proteína-ligante; así como as limitacións inherentes ao extrapolar os resultados teóricos aos experimentais. Neste contexto, continúanse desenvolvendo novas funcións de puntuación para unha mellor estimación da afinidade termodinámica dos sistemas proteína-ligante. Isto abre un horizonte ilimitado de posibilidades para potenciais aplicacións no campo da Química Medicinal e Nanobiotecnoloxía.

Para abordar este tema unha introdución xeral (**Capítulo 1**) estará dedicada aos aspectos relevantes da estrutura e función das proteínas relacionando os seus diferentes

niveis de organización estrutural, seguido dunha definición xeral de receptores diana, así como a definición xeral de "ligante" e a súa clasificación desde o punto de vista farmacolóxico e da Química Medicinal. Posteriormente, trataranse os principais aspectos da avaliación termodinámica da interacción proteína-ligante e os diferentes tipos de interaccións posibles; proporcionando un marco conceptual a un nivel máis teórico. Neste contexto, abordaranse a estabilidade termodinámica dos complexos de axuste xerados en termos de afinidade de unión, e a súa importancia para asegurar a selectividade, especificidade e/ou modulación alostérica dos fármacos nos sitios activos de receptores diana, e ao mesmo tempo abordarase o tema interaccións de unión indesexables (exemplo: fármaco-fármaco ou bloqueo de canles transportadores), comunmente asociadas con efectos secundarios farmacolóxicos ou nanotoxicidad (cando referidos a efectos derivados de nanopartículas ou nanomateriales).

Adicionalmente, os métodos termodinámicos teóricos e experimentais foron amplamente recoñecidos e aplicados ao estudo das interaccións de axuste entre o receptor de proteína e as moléculas de ligante como: fármacos convencionais, péptidos, acedos nucleicos e recentemente produtos derivados da nanotecnoloxía, como os nanotubos de carbono (CNT), fullerenos (C60) e nanomateriales de grafeno (Gr), os cales experimentaron un avance significativo no campo da Química Medicinal e o deseño racional de novos fármacos. Neste contexto, a estabilidade termodinámica dos complexos de axuste proteína-ligante (fármaco-receptor) presentan especial importancia para garantir a selectividade do fármaco nos receptores biolóxicos e, ao mesmo tempo, evitar interaccións indesexables, usualmente asociadas con efectos secundarios farmacolóxicos, mecanismos de toxicidade ou nanotoxicidad (cando se derivan de nanopartículas ou nanomateriales). Os procesos termodinámicos que determinan a natureza do recoñecemento molecular e os eventos de unión entre unha proteína receptora e un ligante dado, son dependentes dos compoñentes entálpicos (ΔH), entrópicos ($-T\Delta S$), así como dos atributos físico-químicos intrínsecos dos ligantes, determinando relacións cuantitativas estrutura-actividade/toxicidade, que poden ser eficientemente preditas a través de modelos QSAR.

Considerando todo o anteriormente exposto, o obxectivo xeral desta tese doutoral está orientado ao estudo e aplicación de métodos teóricos-experimentais para a avaliación de interaccións proteína-ligante en sistemas con potencial relevancia en Biomedicina, Ciencia de Materiais e Saúde Ambiental. Para iso, a segunda sección (**Capítulo 2**) estará dedicada a introducir as principais metodoloxías para o estudo teórico e

experimental da termodinámica de interaccións proteína- ligante. Por tanto, serán abordados os procesos termodinámicos que determinan a natureza do recoñecemento molecular e os eventos de unión proteína e ligante (fármacos). Serán igualmente explicadas as relacións entre parámetros craves como: a enerxía libre de Gibbs (ΔG), entalpía (ΔH), entrópica (-TΔS) e mesmo en relación a conceptos clásicos e emerxentes (escala nanométrica) sobre relacións-cuantitativa-estrutura-actividade e a súa importancia preditiva. Aquí, cabe destacar que ao longo desta tese será particularmente salientada a importancia dos métodos teóricos sobre os experimentais.

Tendo isto vista, prestarase especial atención aos seguintes métodos: i) técnicas de simulación de axuste molecular, ii) cribado virtual baseado en estrutura (VS), iii) enfoques fractales clásicos para resolver problemas de complementariedade proteína-ligante, iv) modelos de resposta a perturbacións baseados en análises de modos normais para estudar propagación de sinais alostéricas de proteínas en presenza e ausencia de ligantes, e v) tamén serán abordados modelos preditivos de relación cuantitativa estrutura- nanotoxicidad (modelos Nano-QSTR).

Os seguintes capítulos conteñen as principais contribucións da presente tese doutoral e resumo todos os procedementos metodolóxicos e achados teóricos máis relevantes obtidos polo Autor.

As tarefas específicas delineadas aquí, estarán orientadas fundamentalmente a investigar teoricamente as interaccións en sistemas proteína-ligante a través de exemplos representativos incluíndo produtos farmacéuticos e nanopartículas de carbono (ou nanomateriales). Especificamente, investigaranse as interaccións con diferentes receptores diana de interese biomédico e saúde ambiental (exemplo, fibrinógeno, canle mitocondrial humano dependente de voltaxe, encima F0F1-ATPasa de *Saccharomyces cerevisiae*, proteína tau biomarcador relevante de Alzheimer e proteína metabolizadora de substratos hidrocarburos aromáticos derivados do petróleo).

Comezando polo **Capítulo 3**, o cal examinará as propiedades estruturais e funcionais responsables da interacción entre os fármacos beta- bloqueadores clásicos (acebutolol e propranolol) co fibrinógeno plasmático como receptor. A continuación, o **Capítulo 4** estará dedicado a investigar a influencia de múltiples modos de unión do fármaco benzotiazol (BTS) interactuando igualmente co fibrinógeno plasmático como receptor diana, aplicándose para este fin, análise estrutural baseada en modelos fractales para estudar propiedades xeométricas de complementariedade en sistemas proteína- ligante no contexto de estudos simulación de axuste. Posteriormente, o **Capítulo 5** abordará

desde un punto de vista mecanístico a problemática das interaccións farmacolóxicas fármaco-fármaco a través do estudo das interaccións entre os axentes betabloqueadores (acebutolol + propranolol) co fibrinógeno. Seguidamente, no **Capítulo 6** analizarase desde unha perspectiva comparativa, a influencia das propiedades estruturais e conformacionales para dous exemplos de penicilinas antibióticos β-lactámicos (dicloxacilina e cloxacilina) interactuando co fibrinógeno.

Finalmente, as seguintes seccións (**Capítulo 7**, **Capítulo 8**, **Capítulo 9** e **Capítulo 10**) tratarán aspectos relevantes sobre a interacción proteína-ligante para o caso de ligantes de interese no campo da nanotecnoloxía e nanobiotecnología. Para este fin, serán estudadas nanopartículas (nanotubos de carbono, fullerenos funcionalizados) e nanomateriales (grafeno) de carbono interactuando cos receptores moleculares anteriormente citados e que posúen especial relevancia para os campos de investigación emerxentes de: Ciencias dos Materiais, Nanomedicina, Nanotoxicología, Diagnóstico Molecular e Saúde Ambiental. Esencialmente, tamén cabe destacar as aplicacións de enfoques teóricos innovadores para mellorar os tradicionalmente usados no terreo de estudo de Ciencia de Materiais e as súas aplicacións biomédicas.

Por último, no **Capítulo 11** presentarase unha discusión integrada como conclusións da tese doutoral destacando as principais contribucións para a investigación das interaccións proteína-ligante, así como, os principais desafíos e perspectivas futuras para continuar avanzando cara a novos paradigmas metodolóxicos para a avaliación racional e eficiente aplicada ao descubrimento de novos fármacos.

Resumen

Comprender la termodinámica de las interacciones proteína-ligante ha sido fundamentalpara el desarrollo y aplicaciones actuales de la Química Medicinal. Desde que en 1894 Emil Fischer definiera las bases del reconocimiento molecular a través de la formulación del "principio de la llave y la cerradura". El conocimiento acerca del ajuste geométrico entre las partes interactuantes (proteína + ligante), constituye un requisito previo el cual determina y ejerce una influencia directa en la afinidad de las interacciones. Sin embargo, la complementariedad geométrica por sí sola no contribuye a la formación de complejosde acoplamiento proteína-ligante estables. En este contexto, las interacciones no covalentes constituyen la fuerza esencial que sustenta la afinidad de unión entre las moléculas interactuantes (proteína + ligante) para un reconocimiento molecular exitoso. Por supuesto, una aplicación rigurosa del principio de "llave y cerradura" de Emil Fisherdeberá obligatoriamente considerar la flexibilidad o rigidez de los mencionados sistemasde acoplamiento o reconocimiento molecular (proteína + ligante). Al mismo tiempo que deberá considerar criterios sobre el grado de complementariedad y balance entre las propiedades hidrofóbicas e hidrofílicas de proteínas y ligantes.

En este sentido, los recientes avances en las técnicas de cristalografía de rayos X integrados con métodos de Biología Estructural y Computacional han llevado a una explosión significativa en el número de estructuras públicamente disponibles bancos de datos de proteínas (del inglés, "*Protein Data Bank*"), la cual recientemente superó las 123,456 estructuras proteicas cristalografiadas.

Por otro lado, la base de datos denominada "*PubChem*" es una de las bases de datos más reconocidas y en continuo crecimiento conteniendo más de 15,5 millones de entradas sobre compuestos puros y mezclas, lo cual permite la obtención de información cristalográfica relevante para el estudio teórico de las propiedades de interacción de los compuestos ligantes, (incluyendo los compuestos derivados de la nanotecnología).

En este sentido, los estudios termodinámicos sobre los eventos de reconocimiento molecular proteína-ligante son fundamentales para identificar potenciales aplicaciones biomédicas (bioactividad) de nuevas moléculas, jugando desde etapas muy tempranas de la investigación preclínica de nuevos fármacos un papel central en el desarrollo exitoso de nuevos principios activos (fármacos).

Paralelamente se ha constatado que el número y diversidad molecular de compuestos (fármacos, nanopartículas, nanomateriales, etc.) ha experimentado actualmente un crecimiento desproporcionado. Este hecho es consecuencia directa del desequilibrio existente entre la velocidad a la que se desarrollan y se comercializan los nuevos ligantes (fármacos) y los cortos tiempos de evaluación rigurosa sobre los riesgos farmacológicos(interacciones farmacológicas) y toxicológicos potenciales para la salud humana e inclusola seguridad ambiental.

El escenario se vuelve aún más complejo si se considera que la mayoría de los estudios mecanísticos de farmacología y toxicología experimental sobre interacción proteína- ligante, suelen ser muy costosos en términos de infraestructura, recursos humanos y tiempo. Igualmente, las consideraciones bioéticas asociadas a la investigación experimental con animales de laboratorio y su posterior sacrificio al final de cada proyecto de investigación de nuevos fármacos han ganado cada vez mayor interés.

En este contexto, las instituciones académicas, empresas farmacéuticas y la industria de la biotecnología han desarrollado novedosas y eficientes estrategias metodológicas para mejorar la evaluación de las interacciones proteína-ligante orientadas al diseño racional de fármacos en Química Medicinal. Dichos estudios se basan principalmente en el uso deherramientas de la Química Computacional y Bioinformática aplicada al estudio termodinámico de interacciones proteína-ligante (fármaco-receptor). Como ejemplos representativos podemos citar las técnicas de: simulación de acoplamiento molecular, cribado virtual, enfoques biofísico-mecanísticos utilizando modelos de perturbación local y modelos predictivos sobre relaciones cuantitativas estructura-actividad-toxicidad (modelos QSAR/QSTR). Sumándose recientemente a la lista, la aplicación de modelos de aprendizaje automático basados en inteligencia artificial y redes neuronales arificiales. Como complemento de los ensayos termodinámicos proteína-ligante, los métodos computacionales “in silico” han contribuído significativamente a proporcionar interpretaciones mecanísticas sobre las interacciones proteína-ligante con una precisión razonable, eficiencia predictiva y potencial para extrapolación de resultados teóricos a experimentales. En este sentido, las simulaciones teóricas utilizando modelos multiescala permiten el estudio de las propiedades termodinámicas de los complejos proteína-ligante,lo que también es muy útil para la predicción de mecanismos de interacción a través de los distintos niveles de organización (atómico, molecular, celular). Otra de las aplicaciones reciente de los

métodos de simulación de acoplamiento molecular es su integración con técnicas de inteligencia artificial orientadas al diseño racional de nuevos fármacos, las cuales se ha utilizado con éxito para generar nuevos candidatos experimentalmente sintetizables y con alta afinidad de unión a biomoléculas (proteínas) con potencial interés biomédico.

Las simulaciones acoplamiento molecular, predicen teóricamente la afinidad de interacción entre proteínas y ligantes (fármacos), y son ampliamente aceptadas en el diseño computacional de nuevos fármacos. Las bases teóricas del proceso de reconocimiento entre un ligante y su receptor se sustentan en la estructura cristalográfica3D y la conformación de mínima energía adoptada por el/los ligante(s) durante la interacción con el receptor diana, siguiendo los principios de la teoría de "ajuste inducido".

En este contexto, las técnicas de simulación de acoplamiento molecular pueden clasificarse de acuerdo a tres categorías fundamentales basadas en criterios de flexibilidad para el receptor y/o ligante: i) acoplamiento rígido, ii) acoplamiento semiflexible y iii) acoplamiento flexible. En el acoplamiento rígido, la estructura de los elementos que interactúan (proteína y ligante) no varía en términos de distancias y ángulos de enlaces, permaneciendo ambos parámetros constantes. El método es relativamente simple y evalúaprincipalmente el grado de coincidencia conformacional, lo que lo hace más adecuado para el estudio de sistemas macromoleculares (proteínas con múltiples domínos, ADN, otros). En el segundo caso el acoplamiento semiflexible, permite variar las propiedades de flexibilidad conformacional de las dos moléculas que interactúan (proteína y/o ligante)dentro de un cierto rango de grados de libertad conformacional, lo cual lo hace más apropiado para el estudio de interacciones entre proteínas y ligantes pequeños. Para este caso la estructura un determinado ligante se puede cambiar libremente mientras que la estructura del receptor permanece rígida o en algunos casos algunos de los residuos de aminoácidos del receptor (proteína) son definidos como rotables para optimizar el costo computacional durante el acoplamiento molecular. En la tercera categoría, el acoplamiento flexible permite el estudio íntegro de las propiedades de flexibilidad conformacional del sistema (es decir, proteína y ligando simultáneamente), los cuales se pueden evaluar libremente considerando todos los grados de libertad rotacional y conformacional. Este último tipo de acoplamiento molecular consume más recursos computacionales, sin embargo, mejora significativamente la precisión y la

sensibilidad de las simulaciones.

Los capítulos que componen la presente tesis doctoral se han estructurado en nueve secciones centrales con sus sub-secciones correspondientes, las cuales describen aspectos clave sobre la termodinámica e interacciones inter-moleculares proteína-ligante. Abarcando desde el enfoque tradicional de la Química Medicinal hasta el campo emergente y en contínuo crecimiento de la Ciencia de Materiales (Nanotecnología). Específicamente, en el contexto de las aplicaciones biomédicas de las nanopartículas y nanomateriales de carbono.

Actualmente, se evidencia un avance significativo en la comprensión de las fuerzas termodinámicas implicadas en las interacciones proteína-ligante. Aquí serán presentados varios ejemplos que abordan la complejidad de las interacciones proteína-ligante; así como las limitaciones inherentes al extrapolar los resultados teóricos a los experimentales. En este contexto, se continúan desarrollando nuevas funciones de puntuación para una mejor estimación de la afinidad termodinámica de los sistemas proteína-ligante. Esto abre un horizonte ilimitado de posibilidades para potenciales aplicaciones en el campo de la Química Medicinal y Nanobiotecnología.

Para abordar este tema una introducción general (**Capítulo 1**) estará dedicada a los aspectos relevantes de la estructura y función de las proteínas relacionando sus diferentes niveles de organización estructural, seguido de una definición general de receptores diana, así como la definición general de "ligante" y su clasificación desde el punto de vista farmacológico y de la Química Medicinal. Posteriormente, se tratarán los principales aspectos de la evaluación termodinámica de la interacción proteína-ligante y los diferentes tipos de interacciones posibles; proporcionando un marco conceptual a un nivel más teórico. En este contexto, se abordarán la estabilidad termodinámica de los complejos de acoplamiento generados en términos de afinidad de unión, y su importancia para asegurar la selectividad, especificidad y/o modulación alostérica de los fármacos en los sitios activos de receptores diana, y al mismo tiempo se abordará el tema interacciones de unión indeseables (ejemplo: fármaco-fármaco o bloqueo de canales transportadores), comúnmente asociadas con efectos secundarios farmacológicos o nanotoxicidad (cuando referidos a efectos derivados de nanopartículas o nanomateriales).

Considerando todo lo anteriormente expuesto, el objetivo general de esta tesis doctoral está orientado al estudio y aplicación de métodos teóricos-experimentales para la evaluación de interacciones proteína-ligante en sistemas con potencial relevancia en

Biomedicina, Ciencia de Materiales y Salud Ambiental. Para ello, la segunda sección (**Capítulo 2**) estará dedicada a introducir las principales metodologías para el estudio teórico y experimental de la termodinámica de interacciones proteína-ligante. Por lo tanto, serán abordados los procesos termodinámicos que determinan la naturaleza del reconocimiento molecular y los eventos de unión proteína y ligante (fármacos). Serán igualmente explicadas las relaciones entre parámetros claves como: la energía libre de Gibbs (ΔG), entalpía (ΔH), entrópica ($-T\Delta S$) e incluso en relación a conceptos clásicos y emergentes (escala nanométrica) sobre relación cuantitativa estructura-actividad y su importancia predictiva. Aquí, cabe destacar que a lo largo de esta tesis será particularmente enfatizada la importancia de los métodos teóricos sobre los experimentales.

Teniendo esto vista, se prestará especial atención a los siguientes métodos: i) técnicas desimulación de acoplamiento molecular, ii) cribado virtual basado en estructura (VS), iii)enfoques fractales clásicos para resolver problemas de complementariedad proteína-ligante, iv) modelos de respuesta a perturbaciones basados en análisis de modos normales para estudiar propagación de señales alostéricas de proteínas en presencia y ausencia de ligantes, y v) también serán abordados modelos predictivos de relación cuantitativa estructura-nanotoxicidad (modelos Nano-QSTR).

Los siguientes capítulos contienen las principales contribuciones de la presente tesis doctoral y resumen todos los procedimientos metodológicos y hallazgos teóricos más relevantes obtenidos por el Autor.

Las tareas específicas delineadas aquí, estarán orientadas fundamentalmente a investigar teóricamente las interacciones en sistemas proteína-ligante a través de ejemplos representativos incluyendo productos farmacéuticos y nanopartículas de carbono (o nanomateriales). Específicamente, se investigarán las interacciones con diferentes receptores diana de interés biomédico y salud ambiental (ejemplo, fibrinógeno, canal mitocondrial humano dependiente de voltaje, enzima F0F1-ATPasa de Saccharomyces cerevisiae, proteína tau biomarcador relevante de Alzheimer y proteína metabolizadora de sustratos hidrocarburos aromáticos derivados del petróleo).

Comenzando por el **Capítulo 3**, el cual examinará las propiedades estructurales y funcionales responsables de la interacción entre los fármacos beta-bloqueadores clásicos(acebutolol y propranolol) con el fibrinógeno plasmático como receptor diana. A continuación, el **Capítulo 4** estará dedicado a investigar la influencia de múltiples modos de unión del fármaco benzotiazol (BTS) interactuando igualmente con el

fibrinógeno plasmático como receptor diana, aplicándose para este fin, análisis estructural basado en modelos fractales para estudiar propiedades geométricas de complementariedad en sistemas proteína-ligante en el contexto de estudios simulación de acoplamiento. Posteriormente, el **Capítulo 5** abordará desde un punto de vista mecanístico la problemática de las interacciones farmacológicas drug-drug a través del estudio de las interacciones entre los agentes beta-bloqueadores (acebutolol + propranolol) con el fibrinógeno. Seguidamente, en el **Capítulo 6** se analizará desde una perspectiva comparativa, la influencia de las propiedades estructurales y conformacionales para dos ejemplos de penicilinas antibióticos β-lactámicos (dicloxacilina y cloxacilina) interactuando con el fibrinógeno.

Finalmente, las siguientes secciones (**Capítulo 7**, **Capítulo 8**, **Capítulo 9** y **Capítulo 10**) tratarán aspectos relevantes sobre la interacción proteína-ligante para el caso de ligantes de interés en el campo de la nanotecnología y nanobiotecnología. Para este fin, serán estudiadas nanopartículas (nanotubos de carbono, fullerenos funcionalizados) y nanomateriales (grafeno) de carbono interactuando con los receptores moleculares anteriormente citados y que poseen especial relevancia para los campos de investigación emergentes de: Ciencias de los Materiales, Nanomedicina, Nanotoxicología, Diagnóstico Molecular y Salud Ambiental. Esencialmente, también cabe destacar las aplicaciones de enfoques teóricos innovadores para mejorar los tradicionalmente usados en el terreno de estudio de Ciencia de Materiales y sus aplicaciones biomédicas.

Por último, en el **Capítulo 11** se presentará una discusión integrada como conclusiones de la tesis doctoral destacando los principales aportes para la investigación de las interacciones proteína-ligante, así como, principales desafíos y perspectivas futuras para continuar avanzando hacia nuevos paradigmas metodológicos para la evaluación racionaly eficiente aplicada al descubrimiento de nuevos fármacos.

Summary

Understanding the thermodynamic of protein-ligand interactions has been fundamental toapplications Medicinal Chemistry, since Emil Fischer formulated the first model for molecular recognition, the so-called “lock-and-key principle”, in 1894. The geometric fit between lock (protein) and key (ligand), is a prerequisite that can significantly influence the affinity between these interacting partners. However, shape alone does not bind molecules together. The non-covalent interactions are the driving forces for this process,and the strength between the lock and key is a function of the underlying bonding interactions, which also depend on the distance of the interactions. Of course, the flexibility or stiffness of the protein and ligand systems must be considered and evaluatedin order to rigorously apply the Emil Fischer “lock and key principle” along with the degree of complementarity of the hydrophobic and polar parts of protein and ligands.

Considerable advances in X-ray crystallography combined with recent progress in structural and computational biology have led to an explosion in the number of structurespublicly available in the *Protein Data Bank* (PDB), which recently exceeded 123,456 crystallographic structures. On the other hand, the *PubChem database* is one of the most recognized and growing databases with more than 15.5 million entries of pure compounds and mixtures, allowing to obtain structural information for the study of ligand properties(including some ligand-based nanoparticles).

In this regard, thermodynamic studies of the molecular recognition events of protein-ligand interactions are the fundamental processes for identifying the potential biomedicalapplications or “bioactivity” of new molecules (or small ligands) that play a central role in the successful of drug discovery at very early stages of the preclinical investigation.

The number and molecular diversity of ligands (i.e., small drugs, nanoparticles, nanomaterials, etc.) is growing disproportionately. This fact, is a direct consequence of the existing imbalance between the rate at which new drug candidates are being developed or manufactured and the amount subject to rigorous pharmacological and toxicological risks assessment evaluation in order to improve human health and environmental safety. The scenario becomes even more complex when considering that most experimental pharmacology and toxicology studies based on protein-ligand

interactions are usually very costly in terms of infrastructure, human resources and time. It is also worth mentioning the ethical aspect associated with experimental research on animals and theirsubsequent sacrifice at the end of a given drug discovery project. In this context, academic institutions, pharmaceutical companies and the biotechnology industry have developed new and rapid strategies to improve the evaluation of protein-ligand interactions for rational drug-design and Medicinal Chemistry applications. These studies are mainly based on the use of thermodynamic chemoinformatics tools such as: molecular docking, structure-based virtual screening, perturbation approaches to study allosteric signal propagation in proteins under unbound and bound state (ligand interactions), and prediction models for quantitative structure-activity-binding-toxicity relationships (QSBR/QSTR models) and, more recently, the application of artificial intelligence-basedmachine learning models.

As a complement to thermodynamic ligand-protein experiments, *in silico* methods can efficiently provide mechanistic interpretations of protein-ligand interactions with reasonable precision. Besides, theoretical simulations with multiscale models allow the study of the thermodynamic properties of the obtained protein-ligand complexes at different scales, which is also very useful for the prediction of binding interaction mechanisms. Moreover, integrated structure-based molecular docking with de novo drugdesign methods is successfully being used to generate synthesizable small molecules withhigh binding affinity, which is another direction of molecular docking methods.

Molecular docking simulations, in which the interaction affinity between proteins and small ligands is theoretically predicted, are widely accepted in computational drug design. The theoretical basis is that the recognition-binding process for the ligand and receptor isbased on 3D spatial shape and energy matching, which is supported by the theory of "induced fit".

In this context, molecular docking techniques can be broadly classified into three categories based on previously established flexibility criteria: i) rigid, ii) semiflexible, and iii) flexible docking. In rigid docking, the structure of the two interacting molecules (protein and ligand) is not changed, and the bond distances and angles remain constant. The method is relatively simple and mainly assesses the degree of conformational match, which makes it more suitable for the study of macromolecular systems (e.g., ligand interacting with multi-domain proteins, and nucleic acid systems). Semi-flexible docking allows the conformational flexibility

properties for both molecules to be varied within a certain range, making it more suitable for studying the interaction between proteins and small ligands. Here, the molecular structure of the small ligand can be freely changed while the macromolecular protein (receptor) remains rigid or retains some of the rotatable amino acid residues to ensure maximum computational efficiency. On the other hand, flexible docking allows the study of the conformational flexibility properties of thesystem (i.e., protein and ligand), which can be freely evaluated. The latter consumes morecomputational resources but improves accuracy and sensitivity.

The chapters of the current doctoral thesis, have been structured into nine core sections with their corresponding subsections describing key aspects of thermodynamic intermolecular interactions, from traditional Medicinal Chemistry based drug discovery to the emerging and growing field of Materials Sciences toward biomedical applicationsof carbon nanomaterials and nanoparticles.

Considerable progress has been made in understanding the thermodynamic forces involved through the study of protein-ligand interactions and the systematic investigation of artificial or *in silico* protein-ligand complexes. Here, we present several examples that address the difficult problem of protein-ligand interactions and the inherent limitations ofextrapolating theoretical results to experimental ones. In this context, new scoring functions are currently being developed for better estimation of both, theoretical and experimental binding affinities. This opens an unlimited horizon of possibilities for potential Medicinal Chemistry and Nanobiomedical applications.

Then, it will start with a general introduction (**Chapter 1**) covering relevant aspects of protein structure and function with the corresponding organizational levels of structure, followed by an overview of the general definition of target receptors, as well as, the general definition of ligands and how they can be classified from a Pharmacology and Medicinal Chemistry point of view. Next, the main aspects of thermodynamic evaluation of ligand-protein and different types of interactions by providing a conceptual framework at a more theoretical level are covered. In this context, the thermodynamic stability of thegenerated docking complexes (i.e., protein-ligand binding affinity) is addressed, which is of key importance to ensure the selectivity, specificity, and/or allosteric modulation of drugs at the biological targets, while avoiding undesirable off-target binding interactions or drug-dug interactions associated with pharmacological side effects or nanotoxicity effects (when derived from nanoparticles or nanomaterials).

Considering all the above, the general aim of this doctoral thesis is directed towards the study and application of theoretical coupled with experimental methods for the evaluation of protein-ligand interactions on relevant systems with potential applications for Biomedicine, Materials, and Environmental Health Sciences. To this end, the second section (**Chapter 2**) will be dedicated to introduce traditional and new methodologies forthe theoretical and experimental study of thermodynamic of protein-ligand interactions. Thus, the thermodynamic processes that determine the nature of molecular recognition and binding events between a protein and a given ligand will be address associated to Gibbs free energy of binding, enthalpy (ΔH), entropic (-TΔS), and also with quantitative structure-activity relationship concepts. Herein, should be noted that the importance of theoretical methods over experimental ones, will be particularly emphasized along this thesis.

Particular attention will be paid to the following methods: i) molecular docking techniques, ii) structure-based virtual screening (VS), iii) classical fractal approaches to solve protein-ligand complementarity problems, iv) perturbation response models combined with elastic normal mode analysis to study allosteric signal propagation of proteins and ligands under unbound and bound states, and v) predictive nano-quantitativestructure-binding-toxicity relationships (nano-QSBR/QSTR models).

The following chapters represent the main contribution of the current doctoral thesis andsummarize all the theoretical and experimental evidences obtained by the Author.

The specific tasks will be oriented to investigate the protein-ligand interactions involving representative systems belonging to pharmaceuticals (or xenobiotic) and carbon nanoparticles (or nanomaterials). Specifically, will be investigated interactions with different protein receptors (e.g., fibrinogen blood plasma protein, human mitochondrial voltage-gated anion-selective channel 1 (hVDAC1), F0F1-ATPase from Saccharomyces cerevisiae (F0F1-TPase), Alzheimer tau protein (Tau protein), and Pseudomonas putida substrate-specific transport protein (TodX)). Then it will begin with **Chapter 3**, which examines the structural and functional properties of the binding interactions between the classical β-blockers acebutolol and propranolol with the bovine blood plasma protein fibrinogen. Next, the **Chapter 4** is devoted to investigating the influence of multiple binding modes of the ligand benzothiazole (BTS) interacting with the fibrinogen protein (E-region) using structural fractal analysis in the context of docking studies. Followed bythe **Chapter 5** which tackle the polypharmacology drug-drug interaction study between two beta-blocker drugs (i.e.,

ligand-ligand binding event for acebutolol plus propranolol) with the fibrinogen protein. Afterward, the **Chapter 6**, which discusses the influence of structural and conformational binding properties from a comparative perspective for two examples of penicillin-β-lactam antibiotics (dicloxacillin and cloxacillin) interacting with the blood plasma protein fibrinogen.

By the other hand, the following sections (**Chapter 7**, **Chapter 8**, **Chapter 9**, and **Chapter 10**) deal with general aspects of new carbon nanoparticles (carbon nanotubes, fullerenes) and ligand-based nanomaterials (graphene) that interact with the cited above key molecular receptors forming docking complexes (i.e., CNT+hVDAC1, CNT+F0F1-ATPase, C60+Tau protein, and Gr+TodX) which could be relevant to the fields of Materials Sciences, Precision Nanomedicine, Nanotoxicology, Molecular Diagnosis and Environmental Health Sciences. Essentially, it should also be highlighted the application of innovative and cutting-edge theoretical approaches to improve traditional ones in the context of Material Sciences applications.

Lastly, an integrated discussion in the **Chapter 11** as concluding remarks by integrating the main contributions of the current doctoral thesis investigating protein-ligand problems, as well as, the main challenges, and future perspectives will be provided to continue advancing in the current state-of-the-art of methodologies toward efficient drug-discovery applications.

PUBLISHED PAPERS

- **Michael González-Durruthy**, Gustavo Scanavachi, Ramón Rial, Zhen Liu, M. Natália D. S. Cordeiro, Rosangela Itri, Juan M. Ruso. Structural and Energetic Evolution of fibrinogen toward to the betablocker Interactions. *International Journal of Biological Macromolecules*, 137, 405-419, 2019. doi.org/10.1016/j.ijbiomac.2019.06.229.

  Journal Impact Factor (JIF) (2019): 5.162

  CiteScore (2019): 8.5

- **Michael González-Durruthy**, Gustavo Scanavachi, Ramón Rial, Zhen Liu, M. Natália D. S. Cordeiro, Rosangela Itri, Juan M. Ruso. Mapping the underlying mechanisms of fibrinogen benzothiazole drug interactions using computational and experimentais approaches, *International Journal of Biological Macromolecules*, 163, 730-744, 2020. doi.org/10.1016/j.ijbiomac.2020.07.044.

  Journal Impact Factor (JIF) (2020): 6.953

  CiteScore (2020): 8.5

- **Michael González-Durruthy**, Riccardo Concu, Laura F. Osmari Vendrame, Ivana Zanella, Juan M. Ruso, M. Natália D.S. Cordeiro. Targeting beta-blocker drug-drug interactions with fibrinogen blood plasma protein: A computational and experimental study, *Molecules*, 25, 5425, 2020. doi:10.3390/molecules25225425.

  Journal Impact Factor (JIF) (2020): 4.411

  CiteScore (2020): 4.1

- **Michael González-Durruthy**, Ramón Rial, M. Natália D.S. Cordeiro, Zhen Liu, Juan M. Ruso. Exploring the conformational binding mechanism of fibrinogen induced by interactions with penicillin β-lactam antibiotic drugs. *Journal of Molecular Liquids*, 324, 114667, 2021. doi.org/10.1016/j.molliq.2020.114667.

  Journal Impact Factor (JIF) (2021): 6.165

  CiteScore (2021): 6.9

- **Michael González-Durruthy**, Amal Kanta Giri, Irina Moreira, Riccardo Concu, André Melo, Juan M. Ruso, M. Natália D. S. Cordeiro. Computational Modeling on Mitochondrial Channel Nanotoxicity. *Nano Today*, 34, 100913, 2020. doi.org/10.1016/j.nantod.2020.100913.

Journal Impact Factor (JIF) (2020): 20.722

CiteScore (2020): 20.9

- **Michael González-Durruthy**, Riccardo Concu, Juan M. Ruso, and M. Natália D.S. Cordeiro. New Mechanistic Insights on Carbon Nanotubes' Nanotoxicity Using Isolated Submitochondrial Particles, Molecular Docking, and Nano-QSTR Approaches. *Biology*, 10, 171, 2021. doi.org/10.3390/biology10030171.

Journal Impact Factor (JIF) (2021): 5.079

CiteScore (2021): 3.48

- Patrícia Viera de Oliveira, Luiza Goulart, Cláudia Lange dos Santos, Jussane Rossato, Solange Binotto Fagan, Ivana Zanella, M. Natália D. S. Cordeiro, Juan M. Ruso, **Michael González-Durruthy**. Computational Modeling of Environmental Co-exposure on Oil-Derived Hydrocarbon Overload by Using Substrate-Specific Transport Protein (*TodX*) with Graphene Nanostructures. *Current Topics in Medicinal Chemistry*, 20, 1-18. 2020. doi.org/10.2174/1568026620666200820145412.

Journal Impact Factor (JIF) (2020): 3.218

CiteScore (2020): 5.8

Declaration of Authorship

Herein, all theoretical and computational procedures, as well as the conceptualization and predictive measurements that make up each chapter of this thesis, were performed by the author.

Table of Contents.

List of Figures.

List of Tables.

List of Abbreviations:

L – *Ligand*
P – *Protein*
***R* –** *Receptor*
PDB – *Protein Data Bank*
P-L – *Protein-Ligand complex*
R-L – *Receptor-Ligand complex*
A – *Acebutolol*
P – *Propranolol*
***NP* –** *Nanoparticle*
NM – *Nanomaterial*
DDI – *Drug-Drug Interactions*
ΔH – *Change in enthalpy*
ΔS – *Change entropy*
R – *Gas constant (1.99 cal* mol^{-1} deg^{-1}*; 8.31 J* mol^{-1} deg^{-1}*)*
BTS – *3-(2-Benzothiazolylthio)-propane sulfonic acid*
CNT – *Carbon nanotubes*
SWCNT – S*ingle-walled carbon nanotubes*
MWCNT – M*ulti-walled carbon nanotubes*
CDs – *Cyclodextrins*
Fib – *Fibrinogen*
ADP – A*denosine 5'-diphosphate*
ATP – A*denosine 5'-triphosphate*
hVDAC1 – Human v*oltage dependent anion channel isoform 1*
SMILES – *Simplified molecular input line entry system*
FEB – F*ree energy of binding*
DOF – *Degree of freedom*
DCNNs – *Deep Convolutional Neural Networks*
ENM – *Elastic network models*
ANM – *Anisotropic network models*
NMR – *Nuclear magnetic resonance spectroscopy*
LPRS maps – *Local perturbation response scanning maps*
MD – *Molecular docking*
DFT – *Density Functional Theory*
HOMO – *Highest occupied molecular orbital*
LUMO – *Lowest unoccupied molecular orbital*
MDS – *Molecular dynamics simulations*
TEM – *Transmission electron microscopy*
FFT– *Fast Fourier transform*
FDs – *Fractal dimensions*
M – *Mandelbrot fractal set*
J – *Julia sets*

FD_D_{BW}, **FD_D_{B+BW}** and **FD_D_{W+BW}** – *Fractal binary black (B) and white (W) pixels*
IFS – Iterated Function System
VS – *Virtual screening*
Nano (QSAR) – *Nanoparticle Quantitative-Structure-Activity-Relationships*
Nano (QSPR) – *Nanoparticle Quantitative-Structure-Property-Relationships*
Nano (QSTR) – *Nanoparticle Quantitative-Structure-Toxicity-Relationships*
DMSO – *Dimethyl Sulfoxide*
SASA – *Solvent accessible surface*
EGTA – E*thylene glycol-bis (β-aminoethyl)-N,N,N′,N′-tetraacetic acid*
RLM – *Rat liver mitochondria*
SMP – S*ubmitochondrial particles*
LMR – *Multiple linear regression*
PLS – *Partial least squares*
ANN – *Artificial neural Networks*
RMSD – *Root-mean-square-desviation*
SAXS – *Small angle X-ray scattering*
ITC – *Isothermal titration calorimetry*

Chapter 1.

1. Introduction

1.1 Protein structure and target receptors. An overview and general definitions.

Proteins are defined as large biomolecules consisting of one or more sequentially arranged amino acid residues (i.e., they can be any of the total 20 amino-acid residues with 20^{100} different permutations) that are chemically linked to each other by peptide bonds and generally form polypeptide chains [1]. Basic elements of protein are carbon, hydrogen, oxygen and nitrogen as major components, while sulfur and phosphorus are minor components [1, 2]. The three-dimensional structure of the protein determines its native conformational state depending on its primary structure. They are responsible for performing a variety of structural and biochemical functions, namely: protein synthesis, degradation processes, metabolic reactions, DNA replication, transport, and response to external and internal stimuli occurring within living cells and organisms. Proteins have different levels of structural organization or structural complexity such as: i) primary, ii) secondary, iii) tertiary and iv) quaternary [1, 2].

- **Primary structure**: this level represents the order of amino acid residues constructed by the linear sequence in the polypeptide chain, including the location of disulfide bonds, if any. The peptide bond can be defined as a covalent bond formed between the α-amino group of an amino acid and the α-carboxyl group of the adjacent amino acids [1, 2].

- **Secondary structure**: represents the steric relationship of amino acid residues that are close to each other and represent a regular folding of regions of the polypeptide chain. The most representative types of secondary folding are the α-helix and the β-fold. In the rod-shaped α-helix, the geometric arrangement of amino acids exhibits a helical conformation. The β-sheet structural arrangement, on the other hand, exhibits a folded, slightly curved sheet in parallel or antiparallel conformation, in the same or opposite direction, depending on the structural sense. Some regions of the polypeptide chain that do not form a regular secondary structure can be recognized by the fact that they exhibit a coil or loop conformation [1, 2].

- **Tertiary structure**: Refers to the overall 3D spatial arrangement and relationship of amino acid residues to each other, forming regions (or domains) of a single polypeptide chain. In this context, the structure is supported by four different types of interactions among the amino acid side chains: Hydrogen bonds, ionic interactions between oppositely charged groups, hydrophobic forces, and disulfide cross-links [1, 2].

- **Quaternary structure**: Here, the 3D spatial arrangement of proteins consists of two or more polypeptide chains held together by non-covalent interaction forces [1, 2].

1.1.1 Receptor. Definition and examples under study

In general, the receptor could be defined as specific macromolecular targets (i.e., proteins, nucleic acids, carbohydrates) to which a particular ligand binds, inducing changes or affecting the biochemical activity of the receptors [3]. In the present thesis, the following examples of protein receptors were studied: i) the E region of bovine fibrinogen blood plasma protein (PDB ID: 1JY2), ii) the human mitochondrial voltage-gated anion-selective channel (PDB ID: 2JK4), iii) F0F1-ATPase from Saccharomyces cerevisiae (PDB ID: 5BPS), iv) human Azheimer tau protein (PDB ID: 5O3L), v) metabolizing substrate-specific transport protein from Pseudomonas putida (PDB ID: 3BRZ). General details are given in further sections.

1.2 General definition of ligands. Classification

1.2.1 General definition of ligand: from a biochemical, theoretical and experimental pharmacological point of view, a ligand (L) can be defined as any chemical structure (from the micro- to the nanoscale) that is able to interact reversibly or irreversibly with a certain specificity and selectivity with a specific biomolecule (e.g. proteins, carbohydrates and nucleic acids) or a target receptor (i.e. a macromolecule binding site) [4]. In the context of protein-ligand binding modeling, the ligand represents a molecule capable of interacting with a given protein (P) to form a thermodynamically stable complex (with P-L, 1:1 ratio) with spontaneous binding free energy (affinity). Moreover, L is considered to be smaller compared to P. On the other hand, non-specific ligand binding can lead to off-target interactions (potential site effects or tocixity) [5]. For clarity, the studied ligands were classified into 4 categories or classes with the corresponding examples, considering the origin, structure and potential applications, as follows:

1.2.2 Ligand class 1. Definition, origin and examples.

- **Pharmacutical drug**: Considering the U.S. Food and Drug Administration (FDA) criteria. A drug corresponds to a substance or chemical entity of natural or synthetic origin with potential biomedical and clinical applications for use in the prevention, diagnosis, or targeted for therapeutic purposes for a specific disease. These types of compounds are usually recognized by an official pharmacopeia or health protection regulatory organization [4, 6].

- **Xenobiotics**: A xenobiotic is a chemical entity that does not occur naturally or is not expected to be present in the organism. This definition is often applied to chemical structures that are rare or absent in nature because they have been chemically synthesized in the laboratory, although they may have potential preclinical relevance. In most cases, however, they are expected to cause potential harm (adverse effects or toxicity) to living organisms and environmental health [4-6].

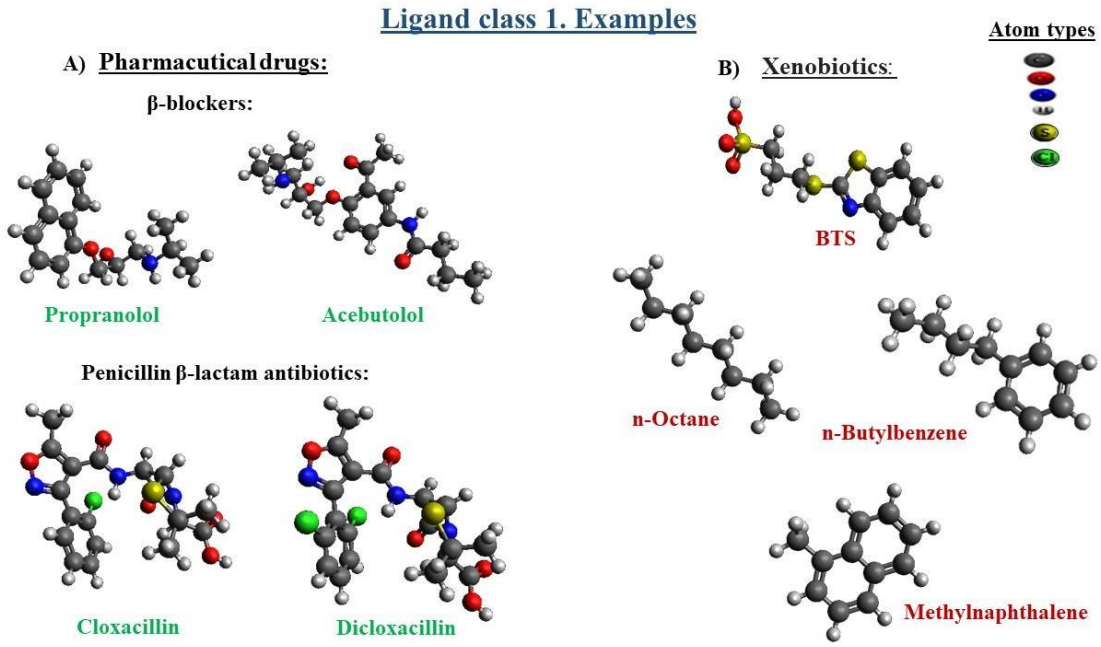


Figure 1. Representative examples of the different ligands studied, belonging to class 1. **A)** pharmaceuticals and **B)** xenobiotics. Further details will be discussed in the next sections.

1.2.3 Ligands class 2. Definition, origin and examples.

- **Nanoparticles (NPs):** This type of ligand can be defined as a natural or synthetically derived chemical entity (or nanoligand-based NP) that has the three dimensions corresponding to the size in the interval of 1-100 nm. Currently, they are efficiently used for diagnostic and therapeutic purposes [7].

- ***Nanomaterials (NMs):*** This type of ligand can be strictly defined as a natural, incidental or manufactured material (organic or inorganic) that has one or more external dimensions in the size range 1-100 nm. Currently, they are efficiently used for chemical catalysis, regenerative medicine and bioremediation of environmental waters [8].

1.2.3.1 General aspects on ligand-based carbon nanoparticles and nanomaterials.

The most commonly used carbon nanoparticles and nanomaterials include the family of carbon nano-allotropes as: carbon nanotubes (CNTs), fullerenes (C60), and graphene (Gr) [9, 10]. These ligands can be created based on the carbon structure from Gr to CNT and C60 with different physicochemical structure and properties. Graphene is an isolated single layer of carbon hexagons consisting of sp^2 hybridized C-C bonds with π-electron clouds. From synthetic technology point of view, thin flakes consisting of few layers of carbon atoms can be very important as mono-layer graphene due to their interesting structural and physical properties and also promising applications in technological fields [9, 10]. The fabrication process of graphene can be classified into five main routes: i) mechanical cleavage of graphite crystals, ii) exfoliation of graphite by its intercalation compound, iii) chemical vapor deposition on various substrate crystals, iv) organic synthesis process, and v) others. The extraordinary chemical and physical properties of Gr lead to a wide range of unusual sensing applications in many fields of science and technology by providing useful solutions to relevant key problems in the world, such as water and food safety and rapid and diagnostic medical analysis [9, 10]. In this regard, the C60 has attracted attention due to its potential application for diagnostic purposes. From a structural point of view, the fullerene C60 ligand nanoparticle corresponds to a truncated icosahedral shape containing 20 hexagons and 12 pentagons [9, 10]. Here, each vertex represents a carbon atom and binds along the polygon edges. In this case, the C60 ligand has a van der Waals diameter of over 1.1 nm and a core-to-core diameter of 0.71 nm. The average bond length is 1.4 Å, and the double bonds between hexagons are

longer than the hexagon-pentagon bonds [9, 10]. On the other hand, in terms of CNT ligands, these carbon nanoparticles correspond to cylindrical structures that have a flat hexagonal network of graphene (carbon allotrope with sp2 hybridization) as the fundamental structural element. CNTs can be naturally occurring, through geological and pyrogenic origin, and are also anthropogenically produced as derived partial products from combustion, unintentional and intentional. The typical CNT structure corresponds to a cylindrical design based on the winding of one or more graphene layers and a length on the order of micrometers [9, 10], so CNTs can occur in the three classical arrangements: single-walled carbon nanotubes (SWCNTs), two cylindrical graphene layers (DWCNTs, double-walled carbon nanotubes), or multi-walled carbon nanotubes (MWCNTs) when composed of two or more concentric graphene layers [9-12]. (See **Figure 2**; below for representative examples of ligand-based carbon NP and NM, which belong to class 2 studied in this thesis.

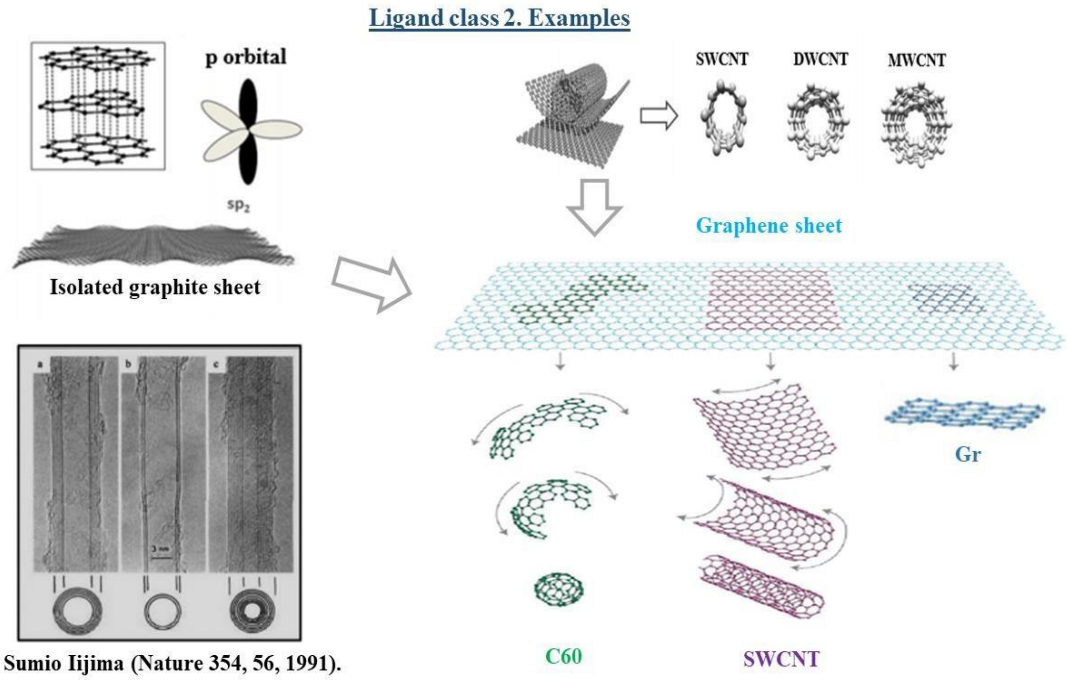


Figure 2. Origin and molecular structure of carbon nanoparticles (C60 and CNTs) and nanomaterials (Gr) belonging to the studied ligand class 2.

From a structural point of view, most of the properties of CNTs (SWCNTs) depend on their diameter (D) and the "chiral angle", also called chirality or helicity (Φ) [13-16]. These two parameters are defined by two non-negative integers called Hamada indices (n, m). Conceptually, a SWCNT can be modeled from a single sheet of Gr rolled to fit two crystallographically equivalent sites of the hexagonal Gr network. Here, the vector C, called

"chiral vector", defines the direction in which the graph sheet is rolled, and the relative position of the two crystallographically equivalent sites is defined by the Hamada indices (n, m) and the unit vectors of the hexagonal lattice a1 and a2 (C = na1 + ma2) [13-16]. This characteristic determines the specific atomic structure of the SWCNT. Thus, a SWCNT is called "armchair" if n = m and Φ=30°, and "zig-zag" if n > 0, m = 0 and Φ = 0°. In all other cases, SWCNTs are called "chiral" when n ≠ m > 0 and 0° ≤ Φ ≤ 30° [13-16]. See **Figure 3**.

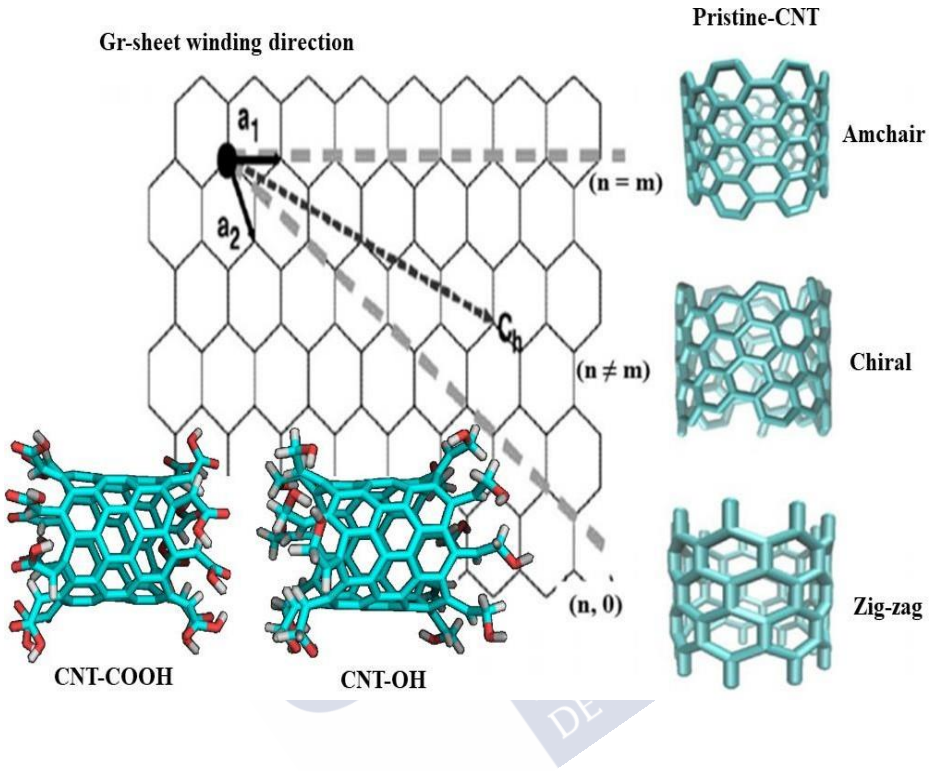


Figure 3. Representation of different models of carbon nanotube ligands (class 2) with different structural geometry by the two Hamada indices (n, m), and type of oxidation (OH, COOH).

Moreover, depending on the Hamada indices (n, m), a given CNT can be classified as metallic (conducting properties) if (n - m) = 3q, where q is an integer, otherwise it is considered as semiconductor. All "armchair" SWCNTs are metallic (q = 0), while "zig-zag" and "chiral" types can be metallic or semiconducting [13-16]. These electrotopological properties of the structure need to be taken into account as they may affect the binding interaction properties (affinity) of CNT ligands with biological components (such as proteins or target receptors) and modulate the reactions at the biochemical level when these CNT structures are used in unmodified or functionalized form (i.e., oxidation with OH groups, e.g. COOH), which may be relevant for the induction of potential nanotoxicity [17-19].

Due to the hydrophobic nature of CNT, C60 and Gr ligands, numerous functionalization methods are currently being developed to improve their biocompatibility for therapeutic purposes in nanomedicine and to reduce their potential nanotoxicity, as well as for environmental safety (mainly for the construction of efficient water filters) [17-19]. Functionalization of CNTs can be achieved by chemical adsorption and/or doping through non-covalent modification, where several functional groups of different complexity can be incorporated in combination with a variety of molecules, including: Peptides, acids, amines, polymers with high detection sensitivity and molecular recognition properties for effective controlled release drug carriers with anticancer, antiviral or antibacterial properties [18]. In particular, the covalent modification of CNTs by oxidative processes often generate hydroxyl (OH) and carboxyl (COOH) groups in the tips and/or walls of CNTs, Gr, and also C60 reactive surface, allows them to express interesting properties such as complexation/chelation of potentially toxic environmental compounds such as aromatic hydrocarbon compounds. In this regard, previous studies show that the oxidation processes of Gr surface significantly enhance the adsorption properties of oxidized Gr. Because the oxidized surface moieties (epoxy, OH and COOH) present in the oxidized Gr reactive surface confer high affinity to strongly interact with polycyclic aromatic hydrocarbons, which usually have high environmental pollution.

1.3 Ligand affinity. Definition

A first formal definition was provided by Irving Langmuir Kenakin in 2004 [4, 20]. It describes affinity as the extent, proportion, or binding interaction strength to which a ligand (L) interacts with the protein receptor (P) at a given concentration to form a thermodynamically stable complex (e.g., P-L complex) with spontaneous free energy. See below the equilibrium reaction and the corresponding **equation 1**.

$$[\boldsymbol{P}] + [\boldsymbol{L}] \underset{\boldsymbol{K_d}}{\overset{\boldsymbol{K_i}}{\rightleftharpoons}} [\boldsymbol{PL}]$$

$$affinity\ (K_i) = \frac{1}{K_d} = \frac{[L][P]}{LP} \quad (\mathbf{1})$$

Where [P], [L] and [P-L] represent the free concentration of protein, ligand and protein-ligand complex, respectively. It is assumed that the components are free to diffuse into the medium

in the equilibrium reaction and that P and L are not changed after dissociation. Here, K_i and K_d are the rate constants of affinity and dissociation of the P-L complex, respectively. Moreover, in the context of pharmacology (i.e., to refer to drug-receptor interaction), fractional occupancy is more commonly used, which also includes these parameters according to **equation 2** below:

$$f_{occupancy} = \frac{[L]}{[L] + [K_d]} \quad (\mathbf{2})$$

In this context, the term fractional occupancy better describes the fraction of the protein receptor that is occupied at a given ligand (or drug) concentration [4, 20]. This means that the binding interaction (effect) of a given drug should depend directly on the fraction of protein receptors occupied.

1.4 Thermodynamics evaluation of ligand-protein interactions.

Theoretical and experimental thermodynamics methods are widely applied to study the docking-binding interactions between protein receptor and ligand molecules, such as conventional pharmaceutical drugs, peptides, nucleic acids, and recently nanobiotechnological compounds, including drug-based carbon nanoparticles (NPs) such as carbon nanotubes (CNTs), fullerenes (C60), and graphene nanomaterials (Gr), which are advancing the Medicinal Chemistry field and drug discovery [21-24]. In this context, the thermodynamic stability of the generated docking complexes (P-L) is of particular importance to ensure drug selectivity at the biological targets while avoiding undesirable off-target binding interactions associated with pharmacological side effects or nanotoxicity effects (when derived from NPs or NMs) [17-24]. In this regard, the thermodynamic processes that determine the nature of molecular recognition and binding events between a protein and a given ligand are directly dependent on the enthalpic (ΔH) and entropic (-TΔS) components and also represent a tight quantitative structure-activity relationship (QSAR) of both interacting partners (ie., P plus L), with spontaneous processes driven by negative values of binding free energy (ΔG), which is of utmost importance for structure-based rational drug design and Material Sciences research for therapeutic purposes [21-24].

In order to address these topics didactically, this chapter has been organized in the following order. First, i) the general principles and thermodynamic concepts of protein-ligand interactions, ii) the types of specific binding interactions, iii) the influence of solvent, crystallographic structure water, and bulk water, iv) the classification of hydrophobic effects,

v) the thermodynamic relationship between protein flexibility and allosteric cooperativity, and finally vi) an overview of the state of the art in methods for evaluating the thermodynamics of protein-ligand interactions; with particular emphasis on the theoretical (i.e., structure-based docking methods) versus experimental methods. The following content will be addressed in order:

1.4.1 General principles on protein-ligand interactions.

1.4.2 Thermodynamics componentes of the free energy of binding.

It is well-known that the general drives the non-covalent spontaneous association between a protein and a given ligand interacting with a negative Gibbs free energy of binding, which is the sum of enthalpic and entropic terms in the system, and it can be expressed according to the following equation 3 [24-28]:

$$\Delta G = \Delta H - T\Delta S \quad (\mathbf{3})$$

Here ΔG is the obtained binding free energy (FEB, kcal/mol), ΔH is enthalpy, ΔS is entropy, and T is temperature. Here, the ΔH contribution to ΔG in the P-L complex represents the strength of the interactions (affinity) between the two components. This specificity for molecular recognition includes van der Waals interactions, electrostatics (Coulomb), hydrogen bonding, ion and halogen interactions, and polarization of chemical interaction groups, among others. In this context, the ΔH contribution in the simplest description represents a measure of the overall dynamics in the system. Moreover, changes in the ΔS are directly related to changes in the translational and rotational degrees of freedom of the system (P-L). In this context, it is well known that changes in the conformational entropy can be favorable, thus also reducing the binding entropy cost [24-28]. On the other hand, the influence of solvation as a solvent reorganization in the immediate vicinity of the relevant binding site of the protein under the conditions of ligand interaction can significantly affect the entropic term of the binding free energy. Moreover, we can define the thermodynamic Gibbs equation 4 as:

$$\Delta G = -RTlnK_i \quad (\mathbf{4})$$

Here R stands for the gas constant, T for the temperature, and K_i for the binding affinity constant. This classical formalism emphasises the relationship between the Gibbs free energy and the relative P-L binding affinity. In addition, two events can be defined: i) the free

solvation energy (ΔG_{sf}) in the unbound state of the interacting components and, analogously, the free solvation energy (ΔG_{sb}) of the ligand-protein complex [27-30]. Since the free energy is independent of the path leading from one particular state of the system to another, it is a state function. In this sense, it is important to note that a similar decomposition analysis can be performed separately for the enthalpic and entropic terms, since these thermodynamic parameters are also state functions [27-30].

Then it is possible to represent the experimental or observable free energy according to **equation 5** as follows:

$$\Delta G_{obs} = \Delta G_i + \Delta G_{sb} + \Delta G_{sf} \quad (\mathbf{5})$$

In this context, the process of P-L binding association can be represented as a typical Born-Haber cycle (see **Figure 4**).

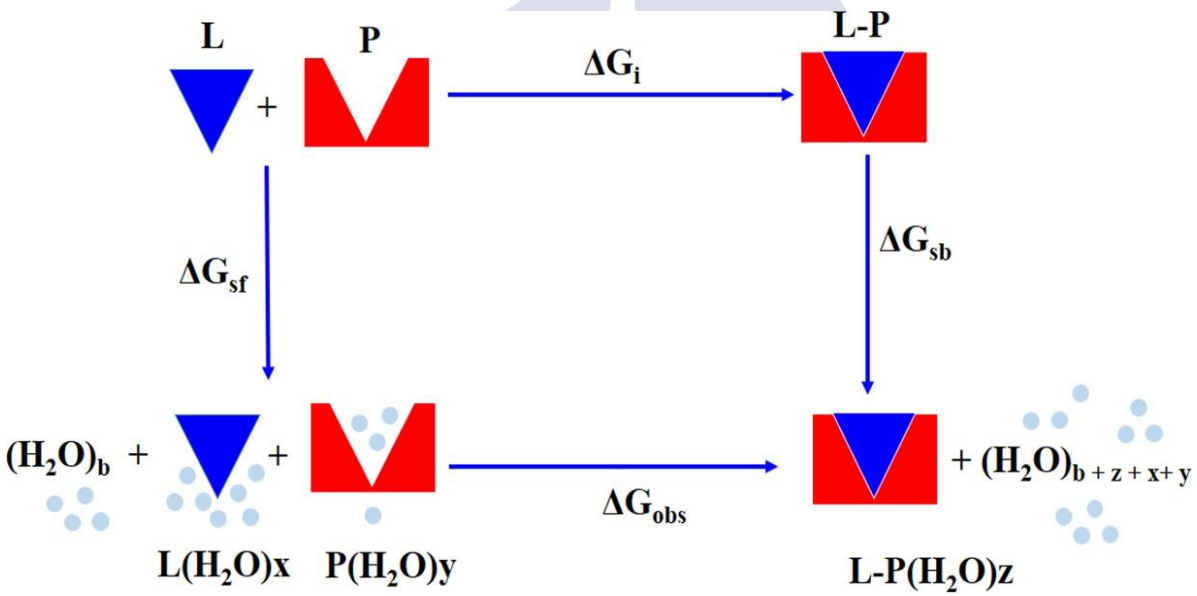


Figure 4. Schematic representation of the Born-Haber cycle for the L-P complex. Here, ΔGobs represents the experimentally observed free energy of binding, ΔGi represents the intrinsic free energy between the ligand (L) and the protein (P), and the free energy of solvation (ΔGsf) for the isolated interacting components (L, P, and waters), and ΔGsb represents the citated complex (L-P) formed. The indices x, y, z and b refer to the number of water molecules involved in the solvation process in the unbound state as: unbound L(x), apo-protein or unbound P (y), the bound P-L(z), and to the bulk solvent (b).

In this context, enthalpic and entropic contributions to a given theoretical or/and experimentally determined binding free energy depend on many properties of the system, such as protonation states, the presence of metal cations, changes in conformational entropy and flexibility properties of both protein and ligands (although these are usually difficult to predict), and others [27-30]. In general, the protein-ligand phenomenon can be addressed by optimising the total free energy of the system under study toward a stable P-L complex (i.e., following the spontaneous thermodynamic process, $\Delta G < 0$ kcal/mol).

1.5 Specific binding interactions

1.5.1 van der Waals interactions

The van der Waals (VdW) interactions are crucial for the structural stabilization of biological molecules and for describing their interactions with ligands. VdW interactions are known to be the weakest of all intermolecular attractions between molecules. In this context, there are two types of van der Waals forces that affect the binding events: (i) attractive van der Waals interactions, which involve two induced dipoles generated by fluctuations in the charge densities between adjacent uncharged atoms of both the protein and the ligand and do not imply covalent bond formation, and (ii) repulsive van der Waals interactions, which occur when the distance between two interacting atoms of the P or L is very small but do not involve induced dipoles. In the latter case, the repulsion is produced by the electron-eléctron repulsion that occurs in the two overlapping electron clouds. The van der Waals interactions between protein and ligand are known to be very weak (0.1-4 kj/mol) compared to the electrostatic or covalent interactions. The van der Waals interactions are often treated as the sum of pairwise interatomic interactions. Moreover, polyatomic VdW interactions (many-body effects), interactions are ignored in most cases because they account for only 5% of the total binding energy. However, Finkelstein (2010) shows that polyatomic VdW interactions can lead to significant changes in the binding free energy in the presence of covalent bonds [31].

1.5.2 Hydrogen bond interactions

Hydrogen bonding interactions (H-bonds) represent non-covalent attractive forces between a hydrogen atom covalently bonded to a specific electronegative moiety (i.e., donor group) and another electronegative atom, such as oxygen or nitrogen (i.e., acceptor group). The H-bonds can be defined as electrostatic dipole-dipole interaction forces. However, H-bonding has certain characteristics or properties of covalent bonding that need to be considered, such as directionality, specificity, and produces interatomic distances shorter than the sum of VdW

radii, which usually involves a limited number of interaction components (P plus L), which can be interpreted as valence type [32-35].

Protein receptors are composed of a large number of H-bond donors and H-bond acceptors, both forming their respective backbones and side chains. In this context, the environment (aqueous solvent, protein-protein network, and lipid bilayers) in which the proteins of interest are embedded also represents a large set of proton donors and acceptors. In this context, H-bonding interactions can occur not only between ligands and proteins or within the protein receptor itself, but also within the surrounding medium such as water molecules, interacting proteins, lipid headgroups or DNA/RNA and so on. As a type of non-covalent interaction, H-bonds are very weak; under physiological conditions, the strength of H-bonds varies from 5-30 kJ/mol and outside the physiological environment, the strength of H-bonds can vary from 2 kJ/mol to even 155 kJ/mol, which is weaker compared to ionic or covalent bonds. However, due to their relative weakness, H-bonds can be rapidly formed and broken during binding events such as protein-ligand interactions, conformational changes of binding sites, and folding processes [32-35]. Therefore, H-bonds can be turned on or off in biochemical systems, with associated energies that are in the range of thermal fluctuations. These events represent one of the key factors favoring protein binding events and biological activity. Another relevant factor in biological systems is the strict geometrical rules followed by H-bonds. In particular, the orientations, angular preferences, and lengths of H-binding make H-binding highly specific. Because of these properties, the role of H-bonds in controlling specific binding interactions in P-L recognition processes is absolutely critical. It is well known that H-bonds (i.e., both intra- and intermolecular) are directly involved in the secondary, tertiary, and quaternary structures of macromolecules (proteins, nucleic acids, and also synthetic polymers). H-bonds play a key role in molecular recognition processes (P plus L) and tune the conformational properties of the macromolecular system (e.g. mechanical strength, bond selectivity and specificity or denaturation processes in proteins and nucleic acids). Hydrogen bonds always impart specificity to a protein-ligand bond recognition process due to their well-aligned geometry and system-specific properties. However, they do not always contribute to high binding free energy [32-35]. Because, H-bonds can change significantly in strength. In this context, stronger H-bonding means higher desolvation in most cases, so the net free energy gain from stronger H-bonding may be greatly affected. As for weak H-bond interactions, the best known donor is the CH group. Thus, despite their weakness, these interactions play a crucial role in the proper conformational stabilization of ligand-protein complexes. As a representative case, it is known that protonated amino acid residues of

histidines (His) can also act as strong CH donors [32-35]. The weak H-bonds and their role in ligand-protein interactions have been systematically studied. [36].

1.5.3 Electrostatic interactions

The electrostatic interactions (Coulomb interactions) are directly involved in the ligand-protein binding processes. This type of intermolecular forces can be broadly divided into three main types: Charge-charge, charge-dipole and dipole-dipole. The classical charge-charge interactions are those that occur between oppositely charged atoms of the two partners, such as functional ligand moieties or protein side chains, such as positively charged (i.e., amine NH_2 or imine groups ($R_2C=NR$) and amino acid residues of Lys, Arg and His) and negatively charged moieties (carboxyl COO^- group, phosphate $(PO_4)^{3-}$ groups, Glu side chain). In the context of protein-ligand interactions (P-L binding), the electrostatic interactions based on the charge dipole of the interacting atoms of P and L have been shown to be an important contributor to the ΔH change associated with the strength of electrostatic interactions that normally occur between ionized amino acid side chains and the dipole belonging to the ligand moiety or water molecule. These dipole moments of the polar side chains forming the amino acid also affect its binding affinity with ligands [37-39].

1.5.4 Hydrophobic interactions

First, it is important to emphasize that the interactions between a particular ligand and the hydrophobic side chains of proteins contribute significantly to the free energy of binding, affinity, and thermodynamic stability of the complexes formed (P-L complexes). Hydrophobic interactions are crucial in rational de novo ligand design because most synthetic drugs must cross pharmacokinetic barriers (i.e., blood-brain barrier, cellular and subcellular membranes, ect) to produce the desired pharmacodynamic effect in a given target protein (P). The hydrophobic of both, protein residues and/or ligands repel each other water and other polar groups. This results in a net attraction of the nonpolar groups of the ligand. Moreover, the apolar and aromatic rings of amino acid residues such as tryptophan (Trp), phenylalanine (Phe), and tyrosine (Tyr) participate in a particular case of hydrophobic interaction, namely π-π-stacking interactions between aromatic groups of the ligand (L) and the protein (P) [40-43]. Several studies indicate that hydrophobic interactions on P-L, when quantified as a function of the amount of hydrophobic surface area of the protein, represent an important structural parameter that is best efficiently correlated with binding free energy or P-L affinity [40-43]. It applies to many and chemically diverse sets of ligands [44], as well as to peptide-

protein interaction phenomena [45]. It should be emphasized that a substantial part of the binding affinity generated by hydrophobic interactions within hydrophobic binding sites is due to the suboptimal solvation of the binding site in the unbound state (apo-P).

1.5.5 Aromatic-aromatic π-π stacking interactions

To understand the physical forces governing ligand-protein interactions, it is important to pay particular attention to the non-covalent interactions between protein side chains: aromatic-aromatic interactions, also called π-π-stacking interactions. The contribution of π-π-stacking interactions to the free energy of binding of P-L complexes has been extensively studied in the context of protein folding, stability (as α-helix and β-sheet stabilisation), molecular recognition, and self-assembly phenomena. Aromatic-aromatic interactions between L- and P-aromatic side chains (i.e., with Phe, Trp, and Tyr) are common in ligand-protein complexes [40]. In particular, the striking steric and electronic properties of the aromatic side chains of Phe, Trp and Tyr lead to these large polarizabilities and quadrupole moments, making oriented geometries under ligand interactions good candidates for geometric studies of ligand binding due to their stereospecific properties. Regarding aromatic-aromatic π-π-stacking interactions. Pioneering studies reported that the phenyl ring centres are separated by a distance of 4.5-7.0 Å [46, 47], and dihedral angles approaching 90° are the predominant ones, with two main types: i) two phenyl ring centres arranged parallel to each other, and ii) a perpendicular orientationthat looks like an edge-to-edge arrangement between the phenyl ring centres. A third type, less common due to its thermodynamic stability, corresponds to iii) parallel-shifted π-π-stacking (i.e., the phenyl ring centres are shifted parallel, here the intermolecular distance for aromatic groups of the protein and the ligand determines the occurrence during P-L binding processes. It is important to note that aromatic-aromatic interactions are not limited to interactions [46, 47]. But also include favourable interactions between heavier halogens and aromatic rings of ligands or proteins.

1.5.6 Halogen bond interactions

A general definition of the concept of halogen-bonding interactions might be similar to the H-bonding interactions discussed earlier. In both types of bonding interactions, interacting atoms of P and L are involved as electron donors and electron acceptors. In H-bonds, an H atom acts as an electron acceptor and forms a non-covalent bond by accepting the electron density from a donor-based electronegative atom [48]. In halogen bonds, a halogen atom corresponds to the donor. This type of interaction has great relevance as small ligands usually contain halogen

atoms to favor their pharmacokinetic bioavailability for medicinal chemistry. The significant role of halogen-bond interactions was largely ignored for many years. The properties of halogen-bond interactions such as specific directionality, sigma holes are generally considered as weak interactions and remain unexplored till date. In some cases, the halogen-bond interactions can compete with H-bonds, providing an attractive alternative for the design of new drugs containing a halogen moiety. In particular, the halogens involved in the halogen-bonding interactions between protein and ligand are chlorine, bromine, iodine, and fluorine. Experimental and computational studies have shown that the four mentioned halogens are able to act as good donors, in the order F > Cl > Br > I, which normally forms the strongest halogen bonds, in terms that the strength increases significantly with the size of the halogen belonging to the ligand (L). From a chemical point of view, the halogens, with the exception of fluorine (F), exhibit unique electronic properties when they interact with aryl or electron-withdrawing alkyl groups. The fluorine halogen exhibits anisotropy of electron density distribution with a positive region (so-called -hole) with electrostatic potential with respect to the carbon-halogen bond [48, 49]. The quantum chemical origin of the -hole can be explained by the formation of a distribution spot of negative charge around the central region of the bond between the carbon atom of a given target protein and the halogen atom of the halogenated ligand, with the outer region being more positive (as a "hole") and having a strong influence on the thermodynamic stability and binding affinity of the formed complexes (halogenated-ligand-protein complex) [48, 49]. In previous reports, replacement of hydrogen with a halogen atom was shown to significantly increase the affinity (200-fold affinity) for a family of complexes belonging to the adenosine kinase inhibitors. Another spectacular evidence was observed for a family of compounds of HIV reverse transcriptase inhibitors with hydrogen substitution by iodine halogen groups, which showed a 300-fold increase in affinity [48, 49]. In this case, the increase in affinity is strongly related to the size of the halogen atom. Moreover, the halogen atoms of the ligand can interact well with the oxygen and carbon atoms belonging to the C=O groups of the proteins, which contributes to the ligand-protein stabilization. Experimental and computational evidence has shown that the commonly used fluoro-halogen from various drugs has a preference for dipolar interactions which is more pronounced than the other halogens [40]. On the other hand, chlorine and other heavy halogens show a great tendency to form multipolar interactions with carbonyl groups of proteins, in which case the C-X halogen bond is parallel to the amide plane rather than orthogonal, leaving the hole behind [40].

1.6 Influence of structural waters and bulk water on binding affinity.

Here it is important to emphasise that each interaction event is capable of displacing water molecules that are part of the biophysical environment of the protein binding site, while simultaneously desolvating the ligand. Since most of these water molecules are only loosely associated with the protein cavities, such displacement can affect the overall solvation envelope around the ligand-protein complex [50-52]. Crystallographic evidence shows that the vast majority of associated water molecules are disordered, mobile, and easily displaced by ligand-specific interactions. Some tightly bound water molecules are usually conserved across multiple crystal structures of ligand-protein complexes [50-52]. Then, the crystallographic water molecules play an important role in the overall biological activity of the protein, mainly in the catalytic processes of enzymes, and can be considered as part of the protein structure [53]. The ligand-protein interactions are usually mediated by structural water buried in the binding sites of the proteins, forming multiple hydrogen bonds. In many cases, the structural water molecules are released from the bulk under ligand interaction. This induces a shift that can dramatically affect the L plus P thermodynamic binding events. In general, it is assumed that water released from a biophysically rigid environment should be entropically favourable. Then the crystallographic structural water molecules can influence their environment not only through the direct interactions (i.e., as a hydrogen bonding network), but also by affecting the dynamic behaviour of their environment as a perturbation of protein flexibility, as suggested by the work performed [50-52]. These flexibility perturbations induced by the water molecules at the Prtein binding sites are highly dependent on the particular structure of the binding site and could likely lead to significant errors in the free energy predictions. The contribution of structural water molecules to the binding events continues to be investigated. This is because the traditional enthalpy-dominated view of ligand-protein interaction largely neglects solvation effects, which strongly influence the thermodynamic profile of a P-L interaction event. This aspect is of key importance since solvation costs are the reason why some ligands, despite being a good fit in a given binding site, fail in experimental validation tests as novel inhibitors of proteins. On the other hand, ligand desolvation itself can sometimes affect P-L binding affinity. In fact, for highly hydrophilic ligands, desolvation costs can be exceptionally high and make an unfavourable contribution to ligand affinity [50-52]. Spectroscopic evidence for the influence of structural water on ligand affinity has shown that: i) water molecules in the first solvation shell surrounding the hydrophobic ligand appear to exhibit high removability, ii) H-bonds placed on hydrophobic surfaces are very weak [50-52], and iii) the removability properties of the

water molecules from the first two solvation shells are significantly different from the bulk water normally bound in buried cavities of the binding site of proteins.

1.7 Thermodynamics of allosteric cooperativity and protein flexibility.

In the context of the study of protein-ligand interactions, it is important to understand the close relationship between protein flexibility and allosteric binding cooperativity, which has been ignored to date and is usually discussed independently of thermodynamic factors. In this context, it is well established that both enthalpic (ΔH) and entropic (-TΔS) contributions are strongly associated. Then, an increase in enthalpy can directly affect the entropy of the system (L-P) under ligand interaction conditions. This is because the presence of mobility constraints (protein stiffening) and/or abnormal flexibility (allosteric perturbations by ligands) can directly affect multiple blocks of adjacent amino acid residues forming an allosteric communication network in the binding sites of a given protein [54-56].

In general, in the presence of external stimuli such as ligand interaction, the binding site cavities undergo a phenomenon known as enthalpy-entropy compensation. In this case, the enthalpy-entropy compensation effects for ligand-protein interactions can be highly cooperative due to the presence of allosteric signals, which means that the associated ΔG is different from the sum of the individual contributions. In this context, allosteric binding cooperativity provides a medium to transmit information about the network of binding residues, to amplify or attenuate a response to changes (i.e., local ligand-induced perturbations) [54-59], and/or to regulate the overall signalling/response pathway. A proper definition of allosteric binding cooperativity could be expressed as a process in which a binding site on a protein molecule (P) is perturbed by an effector (ligand or endogenous allosteric substrate), inducing a structural change at another site and controlling it remotely. Moreover, allosteric binding cooperativity effects can be classified as positive (synergistic or additive) or negative (inhibitory), depending on the intrinsic ability of the interacting ligand to increase or decrease binding affinity for subsequent ligands by altering the flexibility properties of protein's binding sites [54-59]. In contrast, non-cooperative binding effects do not affect the flexibility properties of the protein or the relative affinity for the remaining ligands at subsequent binding sites. In other cases, cooperativity may occur in the absence of any ligand-induced conformational changes in the protein and is driven only by biophysical and physiological changes in intrinsic protein dynamics or the unbound state (Apo-P state) [54-59]. Proteins tend to adaptively compensate for the unfavourable entropic contribution (-TΔS) induced by external or internal stimuli by first increasing their conformational dynamics

in regions distant from the site [54-59]. This section highlights the importance that flexible protein binding pockets may require more flexible ligand units than structurally rigid ones.

In this sense, the traditional study of the ΔH enthalpic term associated with specific interactions and the prevalence of allosteric concepts based on ligand-induced fit theory has led to an overly enthalpic view that ignores the relevance of protein flexibility and its close relationship to low-frequency normal modes of the protein. Following this idea, several studies have used normal mode analysis (i.e., elastic network ENM models) to explain allosteric regulation [54-59]. To this end, ENM models take into account the structural fluctuations of a protein around an equilibrium conformation, where said fluctuations are decomposed into harmonic orthogonal low-frequency modes and, precisely in relation to this, the flexible fluctuations in the network of C-alpha atoms that make up the protein may be sufficient to explain the long-range nature of allosteric communication [54-59]. Currently, these issues are drawing attention to the field of machine-learning allosteric drug discovery using conformational socket flexibility towards de novo rational ligand-based design, which opens new horizons in Medicinal Chemistry.

References Chapter 1

[1] Berg JM, Tymoczko JL, Stryer L. Biochemistry. 5th edition. New York: W H Freeman; 2002. Chapter 3, Protein Structure and Function.

[2] Gromiha MM. Protein Bioinformatics: From sequence to function. Chapter 3. p. 63-105. https://doi.org/10.1016/C2009-0-63223-2.

[3] Lambert DG. Drugs and receptors. Continuing Education in Anaesthesia Critical Care & Pain. 2013, 13, p. 98-102.

[4] Mohammed SS, Prasad SN. An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. Saudi Pharmaceutical Journal. 2017, 25, p 165-175. https://doi.org/10.1016/j.jsps.2016.07.002.

[5] Rudmann DG, On-target and off-target-based toxicologic effects. Toxicol Pathol. 2013, 41, p. 310-314. https://doi.org/10.1177/0192623312464311.

[6] Britannica, The Editors of Encyclopaedia. "Pharmaceutical". *Encyclopedia Britannica*, 9 May. 2018, https://www.britannica.com/technology/pharmaceutical. Accessed 4 August 2021.

[7] Dobson P, King S, and Jarvie H. "Nanoparticle". *Encyclopedia Britannica*, 14 May. 2019, https://www.britannica.com/science/nanoparticle. Accessed 4 August 2021.

[8] European Commission Recommendation of 18 October 2011 on the definition of nanomaterial (2011/696/EU). OJ L 275, 20.10.2011, p. 38-40.

[9] Vasilios G, Jason AP, Jiri T, and Radek Z. Broad Family of Carbon Nanoallotropes: Classification, Chemistry, and Applications of Fullerenes, Carbon Dots, Nanotubes, Graphene, Nanodiamonds, and Combined Superstructures. Chem. Rev. 2015, 115, 11, p. 4744–4822. https://doi.org/10.1021/cr500304f.

[10] Pinzón JR, Villalta-Cerdas A, Echegoyen L. Fullerenes, carbon nanotubes, and graphene for molecular electronics. Top. Curr. Chem. 2012, 312, p. 127-74. https://doi: 10.1007/128_2011_176.

[11] Ravinder K, Sandeep A, Sheng-Joue Y, Vinay G, Pankaj B, and Ajit K. Recent Advances in Carbon Nanomaterials as Electrochemical Biosensors. Journ. Electrochem. Soc. 2020, 167, p. 037555.

[12] Dinadayalane T.C., Leszczynski J. (2012) Fundamental Structural, Electronic, and Chemical Properties of Carbon Nanostructures: Graphene, Fullerenes, Carbon Nanotubes, and Their Derivatives. In: Leszczynski J. (eds) Handbook of Computational Chemistry. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0711-5_22.

[13]. Iijima S. Helical microtubules of graphitic carbon. Nature. 1991. 354, p. 56-58. https://doi.org/10.1038/354056a0.

[14]. Iijima S. Carbon nanotubes: past, present, and future. Physica B: Condensed Matter. 2002, 323, p. 1-5. https://doi.org/10.1016/S0921-4526(02)00869-4.

[15]. Hamada N, Sawada S, Oshiyama A. New one-dimensional conductors: Graphitic microtubes Phys. Rev. Let. 1992, 68, p. 1579. https://doi.org/10.1103/PhysRevLett.68.1579.

[16]. Thomsen C., Telg H., Maultzsch J., and Reich S. Chirality assignments in carbon nanotubes based on resonant Raman scattering. Phys Stat Sol B. 2005, 242, p. 1802- 1806. https://doi.org/10.1002/pssb.200461715.

[17]. Michael G.D et al. QSPR/QSAR-Based Perturbation Theory Approach and Mechanistic Electrochemical Assays on Carbon Nanotubes with Optimal Properties Against Mitochondrial Fenton Reaction Experimentally Induced by Fe^{2+}-overload. Carbon, 2017. 115, p. 312-330. https://doi.org/10.1016/j.carbon.2017.01.002.

[18]. Yang Z., Zhang Y., Yang Y., Sun L., Han D. Pharmacological and toxicological target organelles and safe use of single-walled carbon nanotubes as drug carriers in treating Alzheimer disease. Nanomedicine, 2010. 6, p. 427-41. https://doi.org/10.1016/j.nano.2009.11.007.

[19]. Chistyakov V.A et al. Physical consequences of the mitochondrial targeting of single-walled carbon nanotubes probed computationally. Phys E, 2015. 70, p. 198- 202. https://doi.org/10.1016/j.physe.2015.03.005.

[20]. Kenakin TA. 2004. Pharmacology Primer: Theory, Application, and Methods. Elsevier Academic Press, San Diego & London.

[21]. Bruno OV. Post-Pandemic Drug Discovery and Development: Facing Present and Future Challenges. Front. Drug. Discov. 2021, 1, p. 728469. https://doi.org/10.3389/fddsv.2021.728469.

[22]. Lombardino J, Lowe J. The role of the medicinal chemist in drug discovery-then and now. Nat Rev Drug Discov. 2004, 3, p. 853–862. https://doi.org/10.1038/nrd1523.

[23]. Natalie B et al. The 3Rs as a framework to support a 21st century approach for nanosafety assessment. Nano today, 2017. 12, p. 10-13. https://doi.org/10.1016/j.nantod.2016.06.007.

[24]. Stefano F, Ruth H, Michael E.P, Michel F.S, David S.G, and Arthur J.O. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nat Prot, 2016. 11, p. 905-917. https://doi.org/10.1038/nprot.2016.051.

[25]. Boehm, H.J.; & Klebe, G. (1996). What can we learn from molecular recognition in protein−ligand complexes for the design of new drugs?. Angew. Chem., Int. Ed. 1996, 35, pp. 2588–2614. https://doi.org/10.1002/anie.199625881.

[26]. Dill, K.A. Additivity principles in biochemistry, J. Biol. Chem. 1997, 272, pp. 701-704. https://doi.org/10.1074/jbc.272.2.701.

[27]. Dunitz, J.D. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions., Chem Biol, 1995; 2, pp. 709-12. https://doi.org/10.1016/1074-5521(95)90097-7.

[28]. Gilson MK & Zhou HX. Calculation of protein-ligand binding affinities, Annu Rev Biophys Biomol Struct. 2007, 36, pp. 21-42. https://doi.org/10.1146/annurev.biophys.36.040306.132550.

[29]. Homans, SW. Dynamics and thermodynamics of ligand-protein interactions, Top Curr Chem 2007, 272, pp. 51-82. https://doi.org/10.1007/128_2006_090

[30]. Homans, SW. Probing the binding entropy of ligand-protein interactions by NMR, Chembiochem. 2005, 6, pp. 1585-91. https://doi.org/10.1002/cbic.200500010.

[31]. Finkelstein AV. Average and extreme multi-atom Van der Waals interactions: strong coupling of multi-atom Van der Waals interactions with covalent bonding. Chem Cent J. 2007, pp. 1:21. https://doi.org/10.1186/1752-153X-1-21.

[32]. Deliang C, Numan O, Petri U, Colin F, Sara M.D, Tor C.S. Regulation of protein-ligand binding affinity by hydrogen bond pairing. Sci Adv. 2016, 2, e1501240. https://doi.org/10.1126/sciadv.1501240

[33]. Davis A.M. & Teague S.J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. Angew. Chem., Int. Ed. 1999, 38, pp. 736–749.https://doi.org/10.1002/(SICI)1521-3773(19990315)38:6<736::AID-ANIE736>3.0.CO;2-R.

[34]. Wu M.-Y, Dai D.-Q, & Yan H. PRL-dock: Protein-ligand docking based on hydrogen bond matching and probabilistic relaxation labeling. Proteins: Structure, Function, and Bioinformatics, 2012, 80, pp. 2137–2153. https://doi:10.1002/prot.24104.

[35]. Meyer M, Wilson P, Schomburg D. Hydrogen bonding and molecular surface shape complementarity as a basis for protein docking. J Mol Biol 1996, 264, 199–210. https://doi:10.1006/jmbi.1996.0634.

[36]. Panigrahi S.K., Desiraju G.R. Strong and weak hydrogen bonds in drug-DNA complexes: A statistical analysis. J Biosci. 2007, 32, pp. 677–691. https://doi.org/10.1007/s12038-007-0068-2.

[37]. Bitencourt-Ferreira G, Veit-Acosta M, de Azevedo WF Jr. Electrostatic Energy in Protein-Ligand Complexes. Methods Mol Biol. 2019, 2053, pp. 67-77. https://doi:10.1007/978-1-4939-9752-7_5.

[38]. Aykut E, Monica OC, John FM, Effects of electrostatic interactions on ligand dissociation kinetics. Phys. Rev. E. 2018, 97, p.022405. https://doi:10.1103/PhysRevE.97.022405.

[39]. Alexander D.W, David J.H. Optimization of Protein-Ligand Electrostatic Interactions Using an Alchemical Free-Energy Method. J. Chem. Theory Comput. 2019, 15, pp. 6504–6512. https://doi.org/10.1021/acs.jctc.9b00976.

[40]. Bissantz, C, Kuhn B & Stahl M. A medicinal chemist's guide to molecular interactions. J. Med. Chem. 2010, 53, pp. 5061-5084. https://doi:10.1021/jm100112j.

[41]. Sousa SF, Fernandes PA, Ramos MJ. Protein-ligand docking: current status and future challenges. Proteins: Struct Funct Bioinform. 2006, 65, pp. 15-26. https://doi: 10.1002/prot.21082.

[42]. Gardiner EJ, Willett P, Artymiuk PJ. Graph-theoretic techniques for macromolecular docking. J Chem Inf Comput Sci. 2000, 40, pp. 273–279. https://doi: 10.1021/ci990262o.

[43]. Patil R, Das S, Stanley A, Yadav L, Sudhakar A, Varma AK. Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing. PLoS One. 2010, 5, e12029. https://doi:10.1371/journal.pone.0012029.

[44]. Boehm H.J & Klebe, G. What can we learn from molecular recognition in protein−ligand complexes for the design of new drugs?. Angew. Chem., Int. Ed. 1996, 35, pp. 2588-2614. https://doi.org/10.1002/anie.199625881.

[45]. Vallone B, Miele AE, Vecchini P, Chiancone E, Brunori M. Free energy of burying hydrophobic residues in the interface between protein subunits. Proc Natl Acad Sci U S A. 1998, 95, pp. 6103-6107. https://doi:10.1073/pnas.95.11.6103.

[46]. Zhao Y, Li J, Gu H, Wei D, Xu YC, Fu W, Yu Z. Conformational Preferences of π-π Stacking Between Ligand and Protein, Analysis Derived from Crystal Structure Data Geometric Preference of π-π Interaction. Interdiscip Sci. 2015, 7, pp. 211-20. https://doi: 10.1007/s12539-015-0263-z.

[47]. Brylinski M. Aromatic interactions at the ligand-protein interface: Implications for the development of docking scoring functions. *Chem Biol Drug Des*. 2018;91(2):380-390. https://doi:10.1111/cbdd.13084.

[48]. Suman S, Jake B.B, Rahulsimham V, and Mahesh N. Halogen Interactions in Protein–Ligand Complexes: Implications of Halogen Bonding for Rational Drug Design. J. Chem. Inf. Model. 2013, 53, 11, 2781–2791. https://doi.org/10.1021/ci400257k.

[49]. Clark T, Hennemann M, Murray JS, Politzer P. Halogen bonding: the sigma hole. J. Mol. Model. 2007, 13, pp. 291-296. https://doi:10.1007/s00894-006-0130-2.

[50]. Poornima, C.S.; & Dean, P.M. Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions., J Comput Aided Mol Des. 1995, 9, pp. 500-12. https://doi: 10.1007/BF00124321.

[51]. Poornima, C.S.; & Dean, P.M. Hydration in drug design. 2. Influence of local site surface shape on water binding. J. Comput-Aided Mol. Des. 1995, 9, pp. 513-520. https://doi:10.1007/BF00124322.

[52]. Poornima, C.S.; & Dean, P.M. Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins., J Comput Aided Mol Des. 1995, 9, pp. 521-31. https://doi:10.1007/BF00124323.

[53]. Langhorst U, Loris R, Denisov VP, Doumen J, Roose P, Maes D, Halle B, Steyaert J. Dissection of the structural and functional role of a conserved hydration site in RNase T1. Protein Sci. 1999, 8, pp. 722-30. https://doi:10.1110/ps.8.4.722

[54]. Leach AR. Ligand docking to proteins with discrete side-chain flexibility. J Mol Biol 1994, 235, pp. 345-356. https://doi:10.1016/s0022-2836(05)80038-5.

[55]. Davis IW, Baker D. RosettaLigand docking with full ligand and receptor flexibility. J Mol Biol. 2009, 385, pp. 381-392. https://doi:10.1016/j.jmb.2008.11.010.

[56] Greener, J.G., Sternberg, M.J. AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. BMC Bioinformatics. 2015, **16,** 335. https://doi.org/10.1186/s12859-015-0771-1.

[57] Tama F, Gadea FX, Marques O, Sanejouand YH. Building-block approach for determining low-frequency normal modes of macromolecules, Proteins: Structure, Function, and Bioinformatics. 2000, 41, pp. 1-7. https://doi.org/10.1002/1097-0134(20001001)41:1<1::aid-prot10>3.0.co;2-p

[58] Yang LW, Choong CP. Coarse-grained models reveal functional dynamics - I. elastic network models-theories, comparisons and perspectives, Bioinformatics and Biology Insights. 2008, 2, S460. https://doi.org/10.4137/bbi.s460.

[59] B.H. Lee, S. Seo, M.H. Kim, Y. Kim, S. Jo, M.-k. Choi, H. Lee, J.B. Choi, M.K. Kim. Normal mode-guided transition pathway generation in proteins, PloS one. 2017, 12, https://doi.org/10.1371/journal.pone.0185658.

General objective

The general aim of this doctoral thesis is directed towards the study and application of theoretical coupled with experimental methods for the evaluation of protein-ligand interactions on relevant systems with potential applications for Biomedicine, Materials, and Environmental Health Sciences. It should be noted that the importance of theoretical methods over experimental ones will be particularly emphasised.

Specific objetives

The specific objectives listed below are designed to investigate the protein-ligand interactions of ligands belonging to the previously established ligand classes 1 and 2, such as pharmaceutical drugs (or xenobiotics) and carbon nanoparticles (or nanomaterials), respectively. Specifically, the examples of protein receptors studied were the following: i) bovine fibrinogen blood plasma protein (E-region, PDB ID: 1JY2), ii) human mitochondrial voltage-gated anion-selective channel (hVDAC1, PDB ID: 2JK4), iii) F0F1-ATPase from Saccharomyces cerevisiae (PDB ID: 5BPS), iv) human Azheimer tau protein (PDB ID: 5O3L), v) metabolising substrate-specific transport protein from Pseudomonas putida (TodX, PDB ID: 3BRZ). General details are discussed in a further section.

1. Characterization of the structural and functional properties of the binding interaction between the β-blockers acebutolol and propranolol with fibrinogen protein using molecular docking and normal mode analysis.

2. Determination of the influence of multiple binding modes of the xenobiotic ligand benzothiazole (BTS) interacting on the fibrinogen protein (E-region) using molecular docking simulations coupled with perturbation response scanning maps and structural fractal analysis.

3. Mechanistic investigation of the drug-drug interactions between the beta-blocker drugs (i.e., ligand-ligand ineractions acebutolol plus propranolol) with the fibrinogen protein combining molecular docking and DFT simulations.

4. Theoretical investigation of the structural and functional properties guiding the conformational binding mechanism of fibrinogen through interactions with penicillin-β-lactam antibiotics (dicloxacillin and cloxacillin) using molecular docking simulations

coupled with elastic network models, collective low-frequency normal modes, and perturbation response scanning maps.

5. Evaluate and characterize the binding interactions (potential nanotoxicity) between the zig-zag-single-walled carbon nanotube (z-z-SWCNT) with the ATP-entry-point of the human mitochondrial voltage-dependent anion-selective channel (hVDAC1) using molecular docking, molecular dynamics simulation, and perturbation response scanning maps.

6. Predict and evaluate the mechanism of interaction of a carbon nanotube family (SWCNT-pristine and SWCNT-COOH) with the mitochondrial F0F1-ATPase by using molecular docking, nano-QSBR/QSTR approaches and experimental validation with isolated submitochondrial particles.

7. Investigate the binding interaction properties between Azheimer's disease tau protein and a family of theoretically modeled fullerol-based radiopharmaceutical ligands by integrating molecular docking with DFT approach for diagnostic purposes.

8. Determining the mechanism of interaction of a heterogeneous mixture of aromatic oil hydrocarbons with metabolizing substrate-specific transport protein (TodX) in the presence of graphene nanostructures by combining molecular docking with the DFT approach towards environmental bioremediation strategies.

Chapter 2.

2. Methods

2.1 Methodologies for the study of protein-ligand interactions.

In order to develop accurate techniques for rapid and efficient assessment of ligand-protein interaction phenomena, the use of bioinformatics tools and biophysically inspired approaches was strongly recommended. In this context, molecular docking simulations have been successfully used to rapidly predict the free energy of the binding interaction between a relevant biological target (i.e. protein receptor) of interest and a given ligand (i.e. drug, xenobiotic, nanoparticle and nanomaterials) [1-9]. Molecular docking methods allow modeling, prediction and evaluation of stability associated with the binding affinity of the formed complexes (P-L) [1-9]. To this end, this theoretical approach begins by testing different conformations and orientations of one or a dataset of compounds (structurally related or not) in a presumed binding site (i.e., a site of the protein such as a van der Waal surface cavity) previously defined from the analysis of the three-dimensional structure of the protein target [10-12].

Molecular docking methods for this case usually use linear regression algorithms or mathematical functions ("scoring functions") capable of generating a large number of ligand-protein docking complexes with the corresponding values of binding free energy ($\Delta G \approx$ FEB, expressed in kcal/mol) resulting from the critical balance between favorable and/or unfavorable thermodynamic contributions to affinity [1-12].

The strategy for finding the best crystallographic conformation/orientation of a given ligand in the potential protein receptor binding site must exhaustively explore all possible binding modes of the evaluated ligand, which includes both i) all six translational and rotational degrees of freedom of the ligand and ii) the conformational degrees of freedom of the protein. The process of analyzing the receptor-ligand interaction is not straightforward, as several thermodynamic contributions (i.e., entropics and enthalpics) play a crucial role in P-L affinity when a ligand displaces water molecules from the binding site of the receptor as: (i) hydrophobic van der Waals effects ($\Delta G\_P\text{-}L\ vdW$), hydrogen bonding ($\Delta G\_P\text{-}L\ H\text{-}bonding$), electrostatic interaction between charged atoms ($\Delta G\_P\text{-}L\ electrost$), and finally the internal energy or intramolecular contribution ($\Delta G\_L\ internal$) of the ligand associated with its internal vibrational entropy (ΔS_L) as shown in the following equation:

$$\Delta G_{P-L} = \Delta G_{P-L\_vdW} + \Delta G_{P-L\_H-bond} + \Delta G_{P-L\_electrot} + \Delta G_{L\_int} \quad (\mathbf{1})$$

These thermodynamic-biophysical parameters are discussed in further sections of the thesis. Molecular docking also takes into account the influence of the biophysical environment on the protein in terms of flexibility and/or partial stiffness of the binding sites, co-crystallized ligands from the PDB models of crystallography obtained from *RCSB Protein Data Bank* [13], which should be removed before the docking experiments, the distribution of charges on the ligand under study, as well as the proper selection of specific ligand controls of the simulations, if any, for comparison purposes, as in the case of molecular docking experiments, the interactions of nanoparticles or nanomaterials with specific biological targets (i.e., mitochondrial channels).

2.2 Protein preparation procedures.

In molecular simulation experiments, it is very important to exercise some care to ensure the quality of the computational results and the further possible experimental reproducibility. In this context, it should be noted that the commonly used structure file from the crystallographic model.pdb stored in *RCSB Protein Data Bank* is not suitable for direct use in molecular modeling experiments [13]. This is because some errors originating from the crystallography of the protein need to be corrected beforehand. In this sense, a typical protein.pdb file contains only heavy atoms and may contain structurally embedded co-crystallized ligands, structural water molecules, metal ions and cofactors. In addition, some proteins are multimeric structures (i.e., multidomains) that may need to be reduced to a single domain. Due to the limited X-ray resolution experiments, it could be difficult to distinguish between NH and O. Then, the position of these groups in the protein structure needs to be carefully investigated and verified. In addition, the stored pdb.model may lack structural information in the form of atomic and bond connectivity, which must be assigned along with bond orders and formal charges. These limitations can be overcome by using various modeling tools that automatically correct the structural inconsistencies with little computational effort, such as: i) Schroedinger's Protein Preparation Wizard, ii) Protein_ Prepare Webserver, iii) PROPKA 3.1, iv) PDB2PQR 2.1, and v) Chimera 1.15 software, which also perform combinations of the aforementioned software, among others [14-18]. In this case, several steps are usually required, as shown below. In many cases, however, not all of the steps outlined need to be performed:

1. Import a protein.pdb or a co-crystallized protein ligand structure obtained from Protein Data Bank into the above tools.

2. Find the structural waters you want to keep, then delete all others. Typically, these waters can be identified by the presence of an oxygen atom and usually have no hydrogen atoms attached. Generally, all structural waters (except those bonded or coordinated with metals) are removed. However, structural waters that form a bridge between the protein and the ligand are sometimes retained. If structural waters are retained, then the hydrogen atoms are consistently added to them during protein preparation.
3. Simplification by removing the multimeric complexes, if any. For this purpose, it is necessary to remove redundant structural domains, binding sites and their associated duplicate chains.
4. Correct any serious errors in the protein, metal ions and cofactors, such as incomplete residues, which are very common structural errors, although sometimes they are relatively harmless if they are far from the binding site, while the protein structure with missing residues in the binding site of interest should be repaired obligatorily. Examine the protein structure for metal ions and co-factors in the correct location. If there are bonds to metal ions, it is also necessary to remove the bonds and adjust the formal charge of the atoms bonded to the metal and the metal itself.
5. Correct adjustment of bond orders and formal charge for all cofactors, if necessary.
6. Adjust orientation of misoriented groups, such as amide groups of Asn and Gln.

2.3 Ligand preparation procedures.

Ligand preparation is also a critical step for the success of protein-ligand interaction studies. To this end, we can use the LigPred tools as a robust platform to generate high-quality, all-atom 3D structures starting from 2D or 3D structures for a large number of drug-like molecules used as ligands. To this end, the LigPrep algorithm generates a single 3D structure with high stability-based low energy with the correct chiralities for any input ligand structure. Moreover, LigPrep can also generate a range of ligand structures with different ionization states, tautomers, stereochemistries, and ring conformations, and remove superfluous molecules based on various criteria such as molecular weight or different types of functional groups present in the ligand under study. The process of ligand preparation typically takes over a second per ligand structure output, depending on the type of input structures [14-18].

The general ligand preparation procedure consists of several sequential steps in which multiple conversions and corrections are applied to the ligand structures to generate variations of optimized ligands and remove unwanted features by rigorously adding hydrogen atoms, adjusting the chiralities, examining all chiral atoms in the 3D ligand structures by providing

options to maintain the correct chiralities or to generate additional stereochemical isomers while maintaining the same molecular formula. Tautomerization and ionization states are then adjusted and, if necessary, low-energy ring conformations are generated. The obtained geometries for the prepared structures are then optimized by a short conformational search to relax the ligand structure in 3D dimensions while strongly encouraging the chiral centers to adopt the proper chiralities when the processed structure is highly strained. Finally, a suitable ligand input file is obtained for further analysis (i.e., the study of protein-ligand interaction) [14-18].

2.4 Predicting the ligand-binding site of protein receptors

Correct prediction of binding sites or druggable cavities in proteins is of paramount importance in the context of protein-ligand interaction studies and more generally Medicinal Chemistry. In this context, existing algorithms can treat whole proteins to find suitable binding sites, theoretically functional, and with the correct solvent exposure meet user specifications. A large list of bioinformatics tools has been used for this purpose, such as SiteMap, P2Rank, Pocketome, ezPocket, SiteScore, DeepSite, Q-SiteFinder, ConCavity, CAVER, fpocket, among others [10-12, 19,20]. In general, all of these binding site tools include a scoring function that takes into account the occupancy by hydrophobic groups or by hydrogen bond donors, acceptors, or metal binding capabilities of the ligand to evaluate the propensity of the binding site for ligand binding and to accurately rank possible binding sites to remove irrelevant drug sites [10-12, 19,20].

Currently, the state of the art in binding site prediction has led to multi-criteria algorithms based on artificial intelligence (AI) [2, 10], such as predictive deep learning (DL) tools for predicting possible protein binding sites. In addition, very small cavities within a given binding site can be identified and classified as tunnels (i.e., entangled sub-cavities with significant relevance to de novo drug design). These AI-DL multi-criteria algorithms are based on a 3D convolutional neural network (3D-DCNN) that performs a Delaunay triangulation procedure using the geometric concept of the "alpha-sphere" center, which can detect plausible binding pockets in the protein by summing up all alpha-sphere centers and later narrowing down access to cavities by scanning their van der Waals surface [10-12]. The AI-DL method includes all deep concave regions of the protein that are directly associated with a high probability of ligand binding (i.e., higher drugability) and excludes all convex parts of the protein that are associated with low drugability for ligands [10-12, 19, 20].

Finally, through these different methods, it is possible to set the Cartesian X, Y, and Z coordinates of the box simulation for both the center of the lattice boxand the size of the lattice box like a discretized region of the selected best binding site to perform P-L interaction studies.

2.5 Quality evaluation criteria on protein-ligand docking

One of the most recognized criteria for ensuring the quality of protein-ligand binding studies is the measurement of root-mean-square deviation (R.M.S.D.), which is a docking validation parameter of the correct fit of generated ligand conformations into the binding site of the protein target when R.M.S.D < 2Å is considered a criterion for successful docking [21]. In the context of P-L docking studies, the R.M.S.D values are commonly used to compare a given docked ligand conformation with the reference crystallographic conformation or with other docked ligand conformations generated from the same ligand [21, 22]. Then the R.M.S.D. values are stored according to the following **equation 2**:

$$R.M.S.D = \sqrt{\frac{\sum_{i=1}^{n_L}\sum_{j=1}^{n_P}\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i - z_j\right)^2}{n_L \times n_P}} \quad (2)$$

In the equation, n_L represents the number of atoms of the ligand, while n_P corresponds to the number of atoms of the protein structure being evaluated [21, 22]. Therein, each term included in the double summation represents the distance between a ligand atom i with Cartesian coordinates (x_i, y_i, z_i) interacting with a given protein atom j with Cartesian coordinates (x_i, y_i, z_i).

2.6 Computational tools to study protein-ligand interactions.

2.6.1 AutoDock tools: AutoDock is a commonly used chemobioinformatics program that is one of the most cited software in the scientific community for modeling protein-ligand interactions. In this case, AutoDock uses the Lamarckian Genetic Algorithm (LGA) for predicting/evaluating the best conformation of a studied ligand within the binding site of a protein receptor. The empirical scoring function includes molecular force fields to handle Van der Waals, electrostatics, hydrogen bonding, and desolvation interaction terms [9, 23].

2.6.2 AutoDock Vina: AutoDock Vina software was developed by the same group that developed AutoDock. It includes a new evaluation function that improves the speed and accuracy of docking and is compared to the aforementioned software. The AutoDock Vina has implemented an Iterated Local Search global optimization algorithm for efficient conformational space search [24-27].

2.6.3 DockThor virtual screening: The DockThor- VS web server (available for free at https://www.dockthor.lncc.br) supports both single and structure-based virtual screening (VS) experiments (with up to 5,000 ligands) for registered projects [28]. It also includes a scoring function for soft docking (such as flexible protein-ligand docking). The DockThor scoring function for calculating binding affinities as total intermolecular energy (Etotal, kcal/mol) for the P-L complexes is based on the MMFF94S force field, which represents the sum of thermodynamic terms as: i) the van der Waals bond (with a buffer constant equal to $\delta = 0.35$), ii) electrostatic interaction potentials between the atomic pairs of L and P, iii) the hydrogen bonding interaction energy, iv) the conformational rotation energy, and v) the intramolecular interaction energy term of the studied query ligands. The use of molecular docking simulation coupled with a VS framework is considered a powerful platform for the rational design of new drugs (including NPs and NMs) prior to their massive production, enabling the computational interaction analysis of a large volume of ligand designs with relevant molecular targets (i.e., proteins).

2.6.4 Nano-QSAR tools: This block of computational tools comes from traditional QSAR methods, which have been efficiently adapted to solve problems in nanobiotechnology and materials science. In this context, nanochemoinformatics tools can also be very useful to predict structure-activity/toxicity relationships of carbon nanoparticles and/or nanomaterials (nano-QSAR/nano-QSTR approaches) [29-35]. In particular, the term Nano-Quantitative-Structure-Binding Relationships (nano-QSBR) model is used herein to address predictions of nanoparticle/nanomaterials-protein interactions (i.e., CNT-proteins, C60-proteins, and Gr-protein interactions or nanotoxicity). In the specific case of "nanoparticles or nanomaterials", we can then speak of nano-QSAR (nano-QSTR) or nano-QSBR models by analogy. In general, these nano-QSTR/nano-QSBR models can use the physicochemical properties of NPs or NMs (nano-descriptors) as input to make predictions about their biological activity, nanotoxicity and/or theoretical binding affinities, which in most cases are difficult to rationalize by researchers [29-35]. In this regard, the main assumption of classical

QSAR/QSTR algorithms is that the similar structures (i.e., ligand-based NPs or NMs datasets belonging to the same family) have similar properties. Consequently, small structural changes should be linearly correlated with small changes in the values of their biological properties (nanotoxicity endpoints). Basically, there are a variety of techniques based on statistical data analysis that have been shown to be effective for using nano-QSTR (nano-QSBR models) to predict binding interactions for NPs and NMs with their biological targets. Examples of these techniques are: Linear Discriminant Analysis (LDA), Neural Networks (NN), Random Forest (RF), AI- Deep Learning algorithms, etc [29-35]. Finally, it should be noted that nano-QSTR (nano-QSBR models) should have feature selection capability (FS) to exclude redundant nano-descriptors before nano-QSTR model building[REF28]. In this regard, especially the (n, m)-Hamada indices (which are related to CNT chiral topology and CNT diameter) can become interesting candidates as input variables in nano-QSTR (nano-QSBR modeling). Moreover, the electrotopological and constitutional nanoscriptors derived from the structural information and periodic properties of NPs and NMs can be calculated as 2D/3D nano-desscriptors using Dragon software [29-35].

2.6.5 DFT approaches: The energetic, structural and electronic properties of a given ligand can be efficiently simulated using ab initio DFT calculations [36-39]. Several approaches have been proposed for this purpose. In the present work, we use the DFT simulation implemented with the Spanish Initiative for Electronic Simulations with thousands of atoms (SIESTA code) [36], which is able to perform fully self-consistent calculations by solving the Kohn-Sham equations and using numerical atomic orbitals of the ligands as basis sets [36-39]. In this sense, for all DFT interactions, a cutoff value equal to 200 Ry is used for the lattice integration to represent the charge density of the studied ligands. In this process, the atoms composing the ligand structures are relaxed until the residual forces are less than 0.05 eV/Å for all atoms of the studied ligand. Moreover, the interactions between the nucleus and the valence electrons are described by improved Troullier-Martins pseudopotentials [39], while the molecular orbitals use a localized double zeta plus polarization (DZP) basis set. On the other hand, the correlation energies of the ligand structures can be determined using the local density approximation (LDA) proposed by Perdew and Zunger [38,38]. In fact, the LDA approach has been shown to be more suitable than the generalized gradient approximation (GGA) to study weakly interacting systems such as ligand-target amino acid residues resulting from docking simulations and/or to explore ligand-ligand interactions (i.e., drug-drug interactions with pharmacological relevance for drug discovery) [36-39].

2.6.6 Local perturbation response coupled with fractal approaches.

This theoretical approach allows to evaluate the changes (fluctuations) by scanning or mapping the interatomic distances perturbations induced by a given ligand in the protein residue network (i.e., binding sites). In this regard, the perturbation response induced by the ligands (drugs, xenobiotics, NPs, or NMs) can be evaluated for different amplitudes of shifts from equilibrium in the C(α)-atoms forming the residue network of the protein using normal mode analysis (i.e. i.e. elastic network models - ENM models) are estimated, considering two states: i) the unperturbed state for the unbound protein receptor and ii) the perturbed state describing the protein-ligand complex as a bound state. Here, both states are quantified by the interaction potential (U) as Hooke's potential based on ENM models according to the following equation:

$$U_{p_{(i)}} = \frac{1}{N_a}\sum_{j=1}^{N_a}\left|p_j - u_j\right| \quad (3)$$

Where, $U_{p(i)}$ is the effect of perturbation response in elastic normal mode (i), p_j is the displacement of protein residue (j) in perturbed normal mode (bound state), u_j is the displacement of single protein residue (j) in unperturbed mode (unbound state), and N_a is the number of interacting residues of protein binding site. Then, a 2D matrix is generated based on local perturbation response scanning maps (LPRS maps images) allowing to evaluate allosteric signals propagation such as perturbations in the network of residues (i.e., sensors and effectors of the protein) between the mentioned binding states (i.e., unbound and bound states) which in turn are associated to allosteric cooperative mechanisms and conformational coupling (degree of collectivity) between low-frequency normal modes [40-43].

Next, we can apply fractal analysis by using a combination of two recognized algorithms such as: (i) the fractal Mandelbrot set (cardioid image modeling) and (ii) the box counting method to quantify potential fractal geometric perturbations induced by a given ligand. The key idea behind the proposed fractal approach is the concept of fractal protein, which can be represented as a highly complex fractal shape according to non-Euclidean geometry in both the unbound and bound states, using the irregularity properties of the protein surface and the principles of fractal theory [44-50]. In this context, the main goal of the fractal approach is to theoretically describe the differences in geometric complexity (over multiple scales) and self-similarity patterns from the previously obtained LPRS maps to allow comparison between the two protein states (i.e., unbound and bound states) [44-50]. Additional details are described in another section.

2.7 Experimental validation techniques for protein-ligand interaction

2.7. 1 Isothermal titration calorimetry.

Currently, several experimental tools have been implemented to address relevant problems in protein-ligand binding thermodynamics and, at the same time, as a complement to the validation of the theoretical tools. In this regard, X-ray crystallography of proteins provides important information on enthalpic contributions (i.e., vdW interactions, H-bonds, halogen bonds, electrostatic interactions, etc.). While one usually focuses on describing the properties of static ligand-receptor complexes by providing some information on the enthalpic contribution. However, the temperature factors (B-factors) obtained for heavy atoms (i.e., non-hydrogen atoms of the studied P-L complex) are particularly sensitive to the mean-square shifts of the interacting atoms due to thermal motions that affect the protein-ligand binding dynamics. In this context, isothermal titration calorimetry (ITC) allows the measurement of the heat generated during protein-ligand association [51-54]. To this end, during an ITC experiment, a specific ligand is titrated into a solution containing the protein receptor of interest. The strength or extent of binding is determined by directly measuring the heat exchange between the interacting partners (i.e., L plus P) when heat is absorbed or generated during binding thermodynamics. Therefore, ITC has a central importance as it is the only experimental method where the binding constant (Ki), binding affinity free energy (ΔG), entropy (ΔS), and enthalpy (ΔH) can be determined simultaneously in a single experiment. (Perozzo et al., 2004) Moreover, the ITC experiments performed at different temperatures are commonly used to estimate the heat capacity change (ΔCp) of protein-ligand binding events [51-54]. In the present work, ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal Inc., Northampton, U.S.). To determine the binding isotherms describing the drug-protein interactions. All experimental conditions were carried out at a temperature of 298.15 K and stirring was kept constant at 416 rpm [51-54]. On the other hand, the values obtained from the dilution experiments of pure fibrinogen were systematically removed from the cells containing the drug-protein complexes to ensure that all the heat generated was due to the P-L interaction process only, and blank control experiments were performed using BSA for comparison purposes [51-54].

2.7.2 Small-angle X-ray scattering

Small angle X-ray scattering (SAXS) is a powerful analytical technique that allows the study of nanometric systems (1-100 nm) such as proteins (unbound state) or under the influence of ligands (bound state) forming P-L complexes in situ. For this purpose, SAXS is able to

measure the intensities of X-rays scattered by a sample (i.e., P or P-L samples) as a function of the scattering angle, since the elastic collision between the X-rays and the samples allows to obtain a SAXS profile, even under different experimental conditions (such as different concentrations) [55-57]. In this context, it is important to note that SAXS is considered a recognized method that allows validating theoretical evidence on protein-ligand interactions (such as crystallographic structural or conformational changes), since it is a non-destructive technique that requires a minimal amount of sample for the analysis [55-57]. In the present thesis, SAXS experiments were performed at the SAXS beamline of National Synchrotron Light Laboratory (LNLS) Campinas, Brazil, to validate the theoretical results obtained from molecular docking for the interacting system fibrinogen plus ligand (β-blocker drugs acebutolol and propranolol and for the xenobiotic BTS) at different test concentrations. All measurements were performed at a room temperature of 298 K [55-57].

References Chapter 2.

[1] Waszkowycz B, Clark DE, Gancia E. Outstanding challenges in protein-ligand docking and structure-based virtual screening. Wiley Interdisciplin Rev: Comput Mol Sci .201, 1, pp. 229–259. https://doi.org/10.1002/wcms.18.

[2] Ballester PJ, Mitchell JBO. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. Bioinformatics. 2010, 26, pp. 1169–1175. https://doi.org/10.1093/bioinformatics/btq112.

[3] Fischer D, Lin SL, Wolfson HL, Nussinov R. A geometry-based suite of molecular docking processes. J Mol Biol. 1995, 248, ppp. 459–477. https://doi.org/10.1016/s0022-2836(95)80063-8.

[4] Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ, Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nat. Protoc.2016, 11, p. 905. https://doi.org/10.1038/nprot.2016.051.

[5] Thomsen R, Christensen MH. MolDock: a new technique for high-accuracy molecular docking. J. Med. Chem. 2006, 49, pp. 3315–3321. https://doi.org/10.1021/jm051197e.

[6]. Tao A, Huang Y, Shinohara Y, Caylor ML, Pashikanti S, Xu D. ezCADD: a rapid 2D/3D visualization-enabled web modeling environment for democratizing computer-aided drug design, J. Chem. Inf. Model. 2018, 59, pp. 18–24. https://doi.org/10.1021/acs.jcim.8b00633.

[7]. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. 2011, 51, 10, pp. 2778-2786. https://doi.org/10.1021/ci200227u.

[8]. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. Proteins: Struct Funct Bioinform. 2003, 52, pp. 609-623. https://doi.org/10.1002/prot.10465.

[9]. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem. 1998, 19, pp. 1639-1662. https://doi.org/10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.

[10]. Jiménez J, Doerr S, Martínez-Rosell G, Rose AS, De Fabritiis G. DeepSite: protein-binding site predictor using 3D-convolutional neural networks. Bioinformatics. 2017, 33, pp. 3036-3042. https://doi:10.1093/bioinformatics/btx350.

[11] Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: An open source platform for ligand pocket detection. BMC Bioinformatics. 2009**,** p. 10. https://doi.org/10.1186/1471-2105-10-168.

[12] Feinstein WP, Brylinski M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. J Cheminform. 2015, 7, p. 18. https://doi:10.1186/s13321-015-0067-5.

[13] Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J. The Protein Data Bank and the challenge of structural genomics. Nat Struct Biol. 2000, 7, pp. 957-959. https://doi:10.1038/80734.

[14**]** Schrödinger Release 2021-3**:** Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2021; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2021.

[15] Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des. 2013, 27, p. 221-34. https://doi:10.1007/s10822-013-9644-8.

[16] Beard H, Cholleti A, Pearlman D, Sherman W, Loving KA. Applying Physics-Based Scoring to Calculate Free Energies of Binding for Single Amino Acid Mutations in Protein-Protein Complexes. PLoS ONE. 2013, 8, e82849. https://doi.org/10.1371/journal.pone.0082849.

[17] Martínez-Rosell G, Toni G, De Fabritis G. PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. J. Chem. Inf. Model. 2017, 57, 7, pp. 1511–1516. https://doi.org/10.1021/acs.jcim.7b00190.

[18] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem. 2004, 25, pp. 1605-12. https://doi:10.1002/jcc.20084.

[19] Alasdair TRL, Richard M. Jackson, Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. *Bioinformatics. 2005,* 21, pp. 1908-1916, https://doi.org/10.1093/bioinformatics/bti315.

[20] Krivák R, Hoksza D. P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. J Cheminform. 2018, 10, p. 39. https://doi.org/10.1186/s13321-018-0285-8.

[21] Bordogna A, Pandini A, Bonati L. Predicting the accuracy of protein-ligand docking on homology models. J Comput Chem. 2011, 32, pp. 81-98. https://doi:10.1002/jcc.21601.

[22] Bell EW, Zhang Y. DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *J Cheminform.* 2019, 11**,** 40. https://doi.org/10.1186/s13321-019-0362-7.

[23] Goodsell DS and Olson AJ. Automated Docking of Substrates to Proteins by Simulated Annealing Proteins: Structure, Function and Genetics.1990**,** 8, pp. 195-202. https://doi.org/10.1002/prot.340080302.

[24] Valdés-Tresanco MS, Valdés-Tresanco ME., Valiente PA, *et al.* AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4. Biol Direct, 2020, 15, p. 12. https://doi.org/10.1186/s13062-020-00267-2.

[25] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS and Olson AJ. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. J. Computational Chemistry. 2009, **16,** pp. 2785-91. https://doi.org/10.1002/jcc.21256.

[26]. Stefano F, Ruth H, Michael E.P, Michel F.S, David S.G, and Arthur J.O. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nat Prot, 2016. 11, p. 905-917. https://doi.org/10.1038/nprot.2016.051.

[27]. Quiroga R, Villarreal MA. Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. PLoS One. 2016, 11, e0155183. https://doi:10.1371/journal.pone.0155183.

[28]. Guedes IA, Costa LSC, dos Santos KB, Karl ALM, Rocha GK, Teixeira IM, Galheigo MM, Medeiros V, Krempser E, Custodio FL, Barbosa HJC, Nicolas MF, Dardenne LE. Drug

design and repurposing with DockThor-VS web server focusing on SARS-CoV-2 therapeutic targets and their non-synonym variants. Sci. Rep. 2021, *11*, *p.*1. https://doi:10.1038/s41598-021-84700-0.

[29]. González-Durruthy M, et al. Decrypting strong and weak single-walled carbon nanotubes interactions with mitochondrial voltage-dependent anion channels using molecular docking and perturbation Theory. Sci. Rep. 2017, 7, p. 13271. https://doi:10.1038/s41598-017-13691-8.

[30]. González-Durruthy M, et al. Predicting the binding properties of single walled carbon nanotubes (SWCNT) with an ADP/ATP mitochondrial carrier using molecular docking, chemoinformatics, and Nano-QSBR perturbation theory. RSC Adv. 2016, 6, pp. 58680-58693. https://doi:10.1039/C6RA08883J.

[31]. González-Durruthy M, et al. MitoTarget modeling using ANN-classification models based on fractal SEM nano-descriptors: carbon nanotubes as mitochondrial F0F1-ATPase inhibitors. ACS, J. Chem. Inf. Model. 2019, 59, pp. 86-97. https://doi:10.1021/acs.jcim.8b00631.

[32] González-Durruthy M, et al. Experimental-computational study of carbon nanotubes effects on mitochondrial respiration: In silico nano-QSPR machine learning models based on new Raman spectra transform with Markov-Shannon entropy invariants. ACS J. Chem. Inf. Model. 2017, 57, pp. 1029-1044. https://doi:10.1021/acs.jcim.6b00458.

[33]. González-Durruthy M, et al. QSPR/QSAR-based perturbation theory approach and mechanistic electrochemical assays on carbon nanotubes with optimal properties against mitochondrial Fenton reaction experimentally induced by Fe^{2+}-overload. Carbon, 2017, 115, pp. 312-330. https://doi:10.1016/j.carbon.2017.01.002.

[34]. 1. Santana R. et al. Designing nanoparticle release systems for drug–vitamin cancer co-therapy with multiplicative perturbation-theory machine learning (PTML) models. Nanoscale, 2018, 11, pp. 21811-21823. https://doi:10.1039/C9NR05070.

[35]. González-Durruthy M, *et al*. Mitoprotective activity of oxidized carbon nanotubes against mitochondrial swelling induced in multiple experimental conditions and predictions with new expected-value perturbation theory. RSC Adv., 2015, 5, pp. 103229-103245. https://doi:10.1039/C5RA14435C.

[36] Soler JM, Artacho E, Gale JD, García A, Junquera J, Ordejón P, Sánchez-Portal D. The SIESTA method for *ab initio* order-N materials simulation, Journal of Physics: Condensed Matter. 2002, 14, p. 2745. https://doi:10.1088/0953-8984/14/11/302.

[37] Wang Y, Xu Z, Moe YN. On the performance of local density approximation in describing the adsorption of electron donating/accepting molecules on graphene, Chemical physics. 2012, 406, pp. 78-85. https://doi.org/10.1016/j.chemphys.2012.08.008.

[38] Arrigoni M, Madsen GK. Comparing the performance of LDA and GGA functionals in predicting the lattice thermal conductivity of III-V semiconductor materials in the zinc blende structure: The cases of AlAs and BAs, Computational Materials Science. 2019, 156, pp. 354-360. https://doi.org/10.1016/j.commatsci.2018.10.005.

[39] Troullier N, Martins JL. Efficient pseudopotentials for plane-wave calculations, Physical review B. 1993, 43. https://doi.org/10.1103/PhysRevB.43.1993.

[40] Greener, J.G., Sternberg, M.J. AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. BMC Bioinformatics. 2015, 16**,** 335. https://doi.org/10.1186/s12859-015-0771-1.

[41] Tama F, Gadea FX, Marques O, Sanejouand YH. Building-block approach for determining low-frequency normal modes of macromolecules, Proteins: Structure, Function, and Bioinformatics. 2000, 41, pp. 1-7. https://doi.org/10.1002/1097-0134(20001001)41:1<1::aid-prot10>3.0.co;2-p.

[42] Yang LW, Choong CP. Coarse-grained models reveal functional dynamics - I. elastic network models-theories, comparisons and perspectives, Bioinformatics and Biology Insights. 2008, 2, S460. https://doi.org/10.4137/bbi.s460.

[43] B.H. Lee, S. Seo, M.H. Kim, Y. Kim, S. Jo, M.-k. Choi, H. Lee, J.B. Choi, M.K. Kim. Normal mode-guided transition pathway generation in proteins, PloS one. 2017, 12, e0185658. https://doi.org/10.1371/journal.pone.0185658.

[44] Li J, Du Q, Sun C. An improved box-counting method for image fractal dimension estimation, Pattern Recogn. 2009, 42, pp. 2460-2469. https://doi.org/10.1016/j.patcog.2009.03.001.

[45] Alia M, Samsudin A. New key exchange protocol based on Mandelbrot and Julia fractal set, International journal of computer science and network security. 2007, 7, pp. 302–307.

[46] Y. Fisher, Fractal Image Compression: Theory and Application, Springer Science & Business Media, Place Published, 2012.

[47] Mandelbrot BB. Plane DLA is not self-similar; is it a fractal that becomes increasingly compact as it grows? Physica A: Statistical Mechanics and its Applications. 1992, 191, pp. 95–107. https://doi.org/10.1016/0378-4371(92)90511-N.

[48] Mandelbrot BB. Fractal aspects of the iteration of z→ Λz (1-z) for complex Λ and z, Ann. N. Y. Acad. Sci. 357 (1980) 249–259. https://doi.org/10.1111/j.1749-6632.1980.tb29690.x.

[49] Saupe D. Efficient computation of Julia sets and their fractal dimension, Physica D: Nonlinear Phenomena. 1987, 28, 358–370. https://doi.org/10.1016/0167-2789(87)90024-8.

[50] Nezadal M, Zemeskal O, Buchnicek M. The Box-counting: Critical Study, 4th Conference on Prediction, Synergetic and More, The Faculty of Management, Institute of Information Technologies, Faculty of Technology, Tomas Bata University in Zlin, HarFA software, October, 2001, pp. 25-26.

[51] Damian L. Isothermal titration calorimetry for studying protein-ligand interactions. Methods Mol Biol. 2013, 1008, pp. 103-18. https://doi:10.1007/978-1-62703-398-5_4.

[52] Perozzo R, Folkers & Scapozza L. Thermodynamics of protein-ligand interactions: history, presence, and future aspects, J. Recept. Signal Transduct. Res. 2004, 24, pp. 1-52. https://doi.org/10.1081/rrs-120037896.

[53] Su H, Xu Y. Application of ITC-Based Characterization of Thermodynamic and Kinetic Association of Ligands with Proteins in Drug Design. Front. Pharmacol. 2018, 9, p. 1133. https://doi: 10.3389/fphar.2018.01133.

[54] Damian L. Isothermal Titration Calorimetry for Studying Protein–Ligand Interactions. In: Williams M., Daviter T. (eds) Protein-Ligand Interactions. Methods in Molecular Biology (Methods and Protocols). 2013, vol 1008. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-62703-398-5_4.

[55] Anne T.T and Dmitri I.S. Weak protein-ligand interactions studied by small-angle X-ray scattering. FEBS Journal. 2014, 281. pp. 1974-1987. https://doi.org/10.1111/febs.12772.

[56] Chen PC, Masiewicz P, Perez K, Hennig J. Structure-based screening of binding affinities via small-angle X-ray scattering. IUCrJ. 2020, 7, pp. 644-655. https://doi.org/10.1101/715193.

[57] Dainese E, Sabatucci A, Cozzani I. Small Angle X-Ray Scattering: A Powerful Tool to Analyze Protein Conformation in Solution. Curr Org Chem. 2005, 9, https://doi.org/10.2174/138527205774610877.

Chapter 3.

Michael González-Durruthy, Gustavo Scanavachi, Ramón Rial, Zhen Liu, M. Natália D. S. Cordeiro, Rosangela Itri, Juan M. Ruso. Structural and Energetic Evolution of fibrinogen toward to the betablocker Interactions. *International Journal of Biological Macromolecules*, 137, 405-419, 2019. doi.org/10.1016/j.ijbiomac.2019.06.229.

Journal Impact Factor (JIF) (2019): 5.162.

CiteScore (2019): 8.5

Author Contribution M.G.-D.; Conceptualization, methodology-based on molecular docking, building elastic network-based LPRS maps perturbation response models, evaluation of frequency-collectivity degree relationship with hierarchical cluster analysis and writing—original draft preparation.

Journal authorization:

Structural and energetic evolution of fibrinogen toward to the betablocker interactions

Author: Michael González-Durruthy,Gustavo Scanavachi,Ramón Rial,Zhen Liu,M. Natália D.S. Cordeiro,Rosangela Itri,Juan M. Ruso

Publication: International Journal of Biological Macromolecules

Publisher: Elsevier

Date: 15 September 2019

Chapter 3.

Structural and energetic evolution of fibrinogen toward to the betablocker interactions

Abstract

We present a computational analysis coupled with experimental studies, focusing on the binding-interaction between beta-adrenoreceptor blocking agents (acebutolol and propranolol) with fibrinogen protein (E-region). Herein, computational modeling on structural validation and flexibility properties of fibrinogen E-region showed that the E-region interacting residues, which form the funnel-shaped hydrophobic cavity for ligand-binding, can be efficiently modeled. The obtained free energy of binding (FEB) values for the docking complexes, namely acebutolol/fibrinogen E-region and propranolol/fibrinogen E-region, were very close and amounted to - 6.9 kcal/mol and -6.8 kcal/mol, respectively. They were supported by a high binding-accuracy (R.M.S.D < 2 Å) for the best crystallographic binding-poses in both cases. In this regard, we identify a docking-mechanism of interaction for the propranolol and acebutolol mainly based on non-covalent hydrophobic contacts with the fibrinogen E-region binding-site. Besides, the beta-adrenoreceptor blocking agents are able to induce local perturbations affecting particularly the fibrinogen E-region allosteric residues linked to significant changes in the inter-residue communication and flexibility properties of residue network. In this sense, we show that the key biophysical parameters as frequency and collectivity degree may be compromised in different ways by the interaction with acebutolol and propranolol. Overall the obtained theoretical evidences fit well with previous experimental ones obtained by our group by using isothermal titration calorimetry and small angle X-ray scattering measurements (refer to https://doi.org/10.1016/j.ijbiomac.2019.06.229), allowing to corroborate and validate the proposed computational approaches.

3.1 Introduction

Nowadays it is well known the extraordinary structural and topological richness that proteins can exhibit. The ability of amino acids to fold and organize depending on temperature and solvent moves away infinitely from any human scale. On top of that, proteins have to perform their functions in biological systems surrounded by high

concentrations of different molecules or entities ranging in size from small ions to cells. This crowded environment remains stable by a delicate balance of countless interactions and reactions, mainly electrostatically and entropically driven[1-3]. Fibrinogen, a primary protein in the plasma associated with clotting process[4], has an important role to play in this regard. Not only is a protein engaging for bleeding problems, it is a high molecular protein containing highly diverse structures and multi-domain architecture[5]. Particularly, the N-terminal central nodule so-called "E-region" (~ 35 kDa) plays a decisive role in the binding of the thrombin molecule and fibrin polymerization during the clotting process. By the other hand, from the structural point of view, E-region presents a unique dimeric constitution formed by A_α and B_β chains with symmetric architecture forming a kind of hydrophobic cavity domain or funnel-shaped *(vide infra*).[6]

These characteristics offer to fibrinogen the necessary tools for high surface activity, such as structural flexibility, high charge density or hydrophobicity. It is precisely this heterogeneous and sensitive structure that gives fibrinogen the opportunity to modulate and optimize, in a cooperatively way, its interactions with different molecules and surfaces[7].

Although the topics discussed above are often addressed separately, there is an important point in common**:** fibrinogen-ligand interactions**.** The extensive bibliography on this topic indicates that the affinity between a ligand and a protein is determined by intermolecular hydrogen bonds and ionic interactions, as well as by the shape and charge complementarity of the contact surfaces of both partners [8, 9]. The complete understanding of the structural basis of these systems depends to a large extent on the detailed three-dimensional information on the protein-ligand interactions. Given the nature of the current tasks, the main obstacles that need to be examined urgently are related to the great difficulties to understand the complex interaction mechanisms, at the molecular level, between fibrinogen and different molecules as well as translating those particular results to a macroscopic scale. For this reason, in this work, we examine case studies representatives of distinct structural types of drugs and discuss the lessons learnt from each one. The selected drugs were acebutolol and propranol.

Acebutolol is one of most prescribed beta-selective betablocker for treating hypertension. It has an aryloxipropranolamine chain as non-selective blocker but linked to an aromatic ring containing a polar-4 substituent. The molecule has the ability to self-assembly and the final result is determined by the rigidity and planarity of the aromatic ring and the

hydrocarbon chain attached to it. The attraction of this molecule is that many of the β-adrenoceptor blocking agents have pharmacological effects independent of their β blocking activity, such as the modification of the cell membrane. These effects are known as membrane stabilizing activity and include: non-specific cardiac depression, myocardial conduction velocity depression and local anesthetic activity. One of the routes of administration of acebutolol is intravenous, where it is immediately delivered to the bloodstream to interact with plasma proteins as fibrinogen [10]. Propranolol is a β-adrenergic blocking agent. It is an optically inactive compound, with only the 1-isomer having adrenergic blocking activity. It is the most commonly used drug for hypertension; it is also used for other disorders such as chronic stable angina. Like acebutolol, its pharmacological effects have its origin in the alteration of the properties of the cell membrane[11].

Once exposed both the idea and the systems to analyze, we turn now to detail the methodology. The changes given by the new events introduced under the control of computational techniques has made them indispensable especially at the atomic and molecular scale. In this way, docking procedure has been here employed. It is a *in silico* procedure based on molecular mechanics widely used to predict energies, binding modes, being able to generate a three-dimensional structures of ligands into a target protein in a variety of positions, conformations and orientations[12, 13]. One of the main limitations of the Docking it is the small area of the conformational space that can be explored. Therefore, it is necessary to look for a balance between the computational cost and the scanned search space. In addition to the computational tools, have been also employed experimental techniques as isothermal titration calorimetry, zeta potential and small angle X-ray scattering (SAXS) to study the cited systems. These techniques not only offer an additional route for exploring these systems at another biological scale, but also complement and extend the knowledgement obtained via computational/theoretical resources.

3.2 Material and Methods

3.2.1 Computational approach

3.2.2 Computational modeling-based molecular docking. To tackle the study of interactions between beta-adrenoreceptor blocking agents (propranolol and acebutolol) with fibrinogen protein (E-region) a theoretical mechanistic study based on molecular docking approach was performed. In this regard, the first step consists of preparing the

protein receptor file (*i.e.*, fibrinogen E-region), which was withdrawal from the *RCSB Protein Data Bank* (PDB) x-ray structures [14] like Crystal structure of the central region of bovine fibrinogen (E-region) at 1.4 angstroms resolution with code *PDB ID:* 1JY2, consequently with the specie model used in the experimental section during the thermodynamic study (Bovine plasma fibrinogen). The first stage consists of preparing the bovine fibrinogen (E-region) structure-file (receptor). For this purpose, the E-region PDB structure of bovine fibrinogen was optimized by using the AutoDock Tools 4 software for AutoDock Vina [15,16]. The procedure includes the removal of crystallographic water and all the co-crystallized ligand molecules, if any.

Afterward, the H-atoms were included in the protein structure, according to appropriate hybridization geometry according to the built-in modules to add Gasteiger-Marsili empirical atomic partial-charges and states of protonation of the bovine fibrinogen.pdb x-ray structure [14]. Next, the beta-adrenoreceptor blocking ligands were obtained from *Pubchem Data Base Chemical Structure Search* like: acebutolol hydrochloride molecule (PubChem CID: 1978; MF: $C_{18}H_{28}N_2O_{24}$) and propranolol hydrochloride (PubChem CID: 62882; MF: $C_{16}H_{22}CL_2O_2$). The ligand optimization was carried out by using the MOPAC extension based on geometry optimization [17]. More details on the molecular structure for both betablocker drugs can be found in the **Figure S1 (A** and **B)**.

Then, to study the ligand docking mechanisms with bovine fibrinogen E-region we use an Autodock Vina scoring function developed by Trott et al. (2016). [15,16]. According to the thermodynamic equations 1 and 2 depicted below:

$$\mathrm{FEB_{dock}} \approx \Delta G_{\mathrm{bind}} = \Delta G_{\mathrm{vdW}} + \Delta G_{H-\mathrm{bond}} + \Delta G_{\mathrm{electrost}} + \Delta G_{\mathrm{int}} \quad (1)$$

$$FEB_{dock} \approx \Delta G_{bind} = \Delta G_{vdW} \sum_{lig_{(i)}-fib(E)_{(j)}} \left(\frac{A_{ij}}{{d_{ij}}^{12}} - \frac{B_{ij}}{{d_{ij}}^{6}}\right) +$$

$$\Delta G_{H-bond} \sum_{lig_{(i)}-fib(E)_{(j)}} E(t) \left(\frac{C_{ij}}{{d_{ij}}^{12}} - \frac{D_{ij}}{{d_{ij}}^{10}}\right) + \Delta G_{elec} \sum_{lig_{(i)}-fib(E)_{(j)}} \frac{q_i q_j}{\epsilon(d_{ij}) d_{ij}} +$$

$$\Delta G_{vdw} \sum_{lig_{(ii)}} \left(\frac{A_{ii}}{{d_{ii}}^{12}} - \frac{B_{ii}}{{d_{ii}}^{6}}\right) + \sum_{lig_{(ii)}} E(t) \times \left(\frac{C_{ii}}{{d_{ii}}^{12}} - \frac{D_{ii}}{{d_{ii}}^{10}}\right) + \sum_{lig_{(ii)}} \frac{q_i q_i}{4(d_{ii}) d_{ii}} +$$

$$\sum_{lig_{(ii)}} \gamma_k \left(1 + cos(\varpi_k \theta_k - \theta_{0k})\right) \quad (2)$$

Herein, ΔG= - RT (ln K_i), R (gas constant) = 1.98 cal·(mol*k)$^{-1}$, and K_i represents the theoretic inhibition constant (at T = 298.15 K). The first term of a 12-6/Lennard-Jones

interaction potential (with 0.5 Å smoothing) that evaluates the van der Waals interactions as A_{ij}/d_{ij}^2 (attractive Guassian function) and B_{ij}/d_{ij}^6 (repulsive or hyperbolics interaction) as canonical Lennard-Jones interactions (ligands-fibrinogen)[24], the Gaussian term is negative and the parabolic positive, d_{ij} is the distance of surface calculated as $d_{ij} = r_{ij} – R_i - R_j$, where the interatomic distance is r_{ij}. R_i and R_j are the radii of the atoms in the pair of interaction of ligands(*i*)-fibrinogen (*j*) atoms[15,16]. The second term is the pair-consisting of a donor and acceptor hydrogen bond (H-bond) as a directional 12-10 H-bond potential term such as C_{ij}/d_{ij}^{12} and D_{ij}/d_{ij}^{10} with Goodford directionality, where $E(t)$ is an angular weight factor which represents the H-bonds directionality and d_{ij} follows the criteria mentioned above. The third term represents the Coulomb electrostatic potential stored in the formed complex (ligands-fibrinogen) $_{ij}$ of N charges (q_i, q_j) of charged atoms pairs of of ligands (propranolol and acebutolol)$_{(i)}$ and fibrinogen$_{(j)}$. In this concern, appropriated Gasteiger partial-atomic charges of the fibrinogen was assigned. Herein, the interatomic distance (d_{ij}) is between the point charges as the positions reference of interaction based on distance-dependent dielectric constant. In the present study each Amber force field term was parameterized with default options for the ligands by summing up individual molecular mechanic contributions as: intra-molecular contributions, aromaticity criterion and the set number of active torsions moving of each ligand following the general steps on preparing the ligand from docking modeling[18-20]. The fourth term ($\Delta G_{internal}$) was applied to validate the internal steric energy of each ligands including dispersion-repulsion energy and a torsional-energies through the sum of the default force field parameters (*i.e.*, ligand conformation-independent parameters of docking scoring function)[15-20].

Following that, the electrostatic terms were considered, and the partial atomic charges of ligands were assigned with the Gasteiger-Huckel algorithm after the addition of polar and non-polar hydrogen atoms. Moreover, the conformational search space of the ligands was experimentally-validated with Autodock default options including a randomized large training dataset for ligands properties with 50 genetic algorithm runs, and 25 million evaluations in each, and all the structural-parameters to predict the best crystallographic docking pose and orientation of the ligands following the coordinate systems of the receptor (fibrinogen). The best root-mean-square deviation (*RMSD*) defined by (equation 3) was stored as a relevant criterion of correct crystallographic binding-pose accuracy for atomic positions ($P_{i\text{-}ligand}$, $P_{j\text{-}fibrinogen}$) below 2Å. [15-21]

$$RMSD(p_{(i)}, p_{(j)}) = \sqrt{\frac{\sum_n (atom_{(i)} - atom_{(j)})^2}{n}} \quad (3)$$

It is important to note that the dimensionality of docking is based on degree of freedom (DOF) of each ligand: ligand atom position/translation (x_i, y_i, z_i = 3), ligand orientation/quarternion ($q(x_i)$, $q(y_i)$, $q(z_i)$, $q(w_i)$ = 4), ligand rotable bonds/torsion (tor_1, tor_2, …, $tor_n = N_{tor}$) and total dimensionality of ligands (total DOF = 3 + 4 + n) which have a significant relevance in the FEB_{dock} based on the intra-molecular contributions.

After that, to validate the 3D x-ray crystallographic structure of bovine fibrinogen E-region model Ramachandran plot and Ramachandran plot quality assessment were executed [29]. In order to avoid obtaining false positives in the generated flexible-docking complexes. This algorithm allows to verify the absence of restricted-flexibility for all residues from bovine fibrinogen E-region based on allowed torsion values of Ψ (*Psi*) versus Φ (*Phi*). For this instance, the all the ligands were established as rigid molecules[22]. The free energy of binding (FEB) of the formed docking complexes were calculated based on the score function which approximates the chemical potentials (ΔG_{bind}).

Before the docking simulations the fibrinogen E-region binding sites were predicted through DeepSite Convolutional Neural Network (DCNNs)[22-25]. This step was carried out as delimiting the access to fibrinogen-cavities, potentially at the van der Waals surface that are likely to bind to the ligands (propranolol and acebutolol). To this end, DeepSite considers all the crystallographic descriptors related to the fibrinogen. The DCNNs algorithm was validate using deep-learning library with an extensive test set with > 6500 proteins[23]. The volumetric maps generated were used to establish the Cartesian coordinates of the box simulations[24] according to fibrinogen E-region grid box size with dimensions of X= 64Å, Y= 52Å, Z= 52Å and fibrinogen E-region grid box center X= 13.1Å, Y= -2.0Å, Z= 12.5Å. For this instance, an exhaustiveness (docking accuracy) option set to 100 was used[25,26].

Lastly, the affinity FEB values (ΔG_{bind}) for the best docking poses of the formed complexes was obtaining like the sum of the individual molecular-mechanics terms of chemical potentials (ΔG): van der Waals interactions (ΔG_{vdW}), hydrogen bond ($\Delta G_{H\text{-}bond}$), electrostatic interactions ($\Delta G_{electrost}$), and intramolecular ligands interactions ($\Delta G_{internal}$) as depicted above (equation 2)[21-26].

The affinity FEB values results are classified-like energetically-unfavorable when the free energy of protein-ligand complexes ≥ 0 kcal/mol, indicating either extremely low or complete absence of affinity; otherwise are classified-like medium to high docking affinity.

3.2.3 3D Lig-Plot diagrams. To determinate the relevant intermolecular interaction between ligands (propranolol and acebutolol) with the bovine fibrinogen the corresponding 3D Lig-Plot diagrams were carried out. To tackle this objective, a software named ezLigPlot was used [25,26]. This software evaluates the non-covalent intermolecular interactions present in a given protein-ligand complex. Then, Lig-Plot automatically generates schematic 3D interaction diagrams like (hydrophobic, H-bond, cation-Π, and Π-Π stacking) with its corresponding interatomic distances (d_{ij}) for all the docking complexes[25-29].

3.2.4 Perturbation response analysis-based elastic network models. This theoretical approach evaluates the degree of change (perturbation) induced by a given ligands (propranolol and acebutolol) in the residues network (fibrinogen E-region) by describing the interaction potential (*U*) for a receptor-ligand complex (R-L) as Hookean potential (or 'springs') based on elastic normal mode analysis[38-40]. Which will be described in detail in the next section for the fibrinogen-ligand interactions. Herein, the ligand perturbation response was estimated by averaging across all the fibrinogen binding-residues for the different amplitude between the perturbed and the unperturbed state of the displacements from equilibrium using anisotropic vibrational modes analysis[30-35].

3.3 Results and discussion

To tackle with precision relevant interactions of fibrinogen with ligand, the knowledge of its structure and physiological function must be treated jointly. In this regard, high-crystallographic resolution structure assays of the N-terminal central E-region of bovine fibrinogen reveals a singular dimeric with symmetric architecture[36-41]. The two fragments of the fibrinogen molecule bond tightly at the middle forming disulfide linkages (α-α crosslinking) and noncovalent hydrophobic contacts in a central nodule. The E-region is formed by A_α and B_β chains which form a kind of funnel-shaped or hydrophobic cavity domain with an unusual and tangled structure; herein, is located in the two outer sides the binding sites for thrombin[36-41]. On the opposite side, the N-

terminal Bβ-γ chains fold into an individually separated domain. Where, a unique pair of adjacent disulfide-bonds locally constrain the two Bβ-γ chains (γ-γ crosslinking) allow to adopt several conformations. This Bβ-γ/Bβ-γ dimeric domain from E-region presents an asymmetry architecture that ensuring the known protofibrins supercoiling in the fibrinogen molecule during critical stages of the blood coagulation process[41]. For all the above mentioned, the proposed computational modeling is focused on the study of the interaction of ligands (propranolol and acebutolol) with the central E-region of the fibrinogen molecule. According to this idea, the general information on fibrinogen E-region molecular structure is provided in the **Figure 1**.

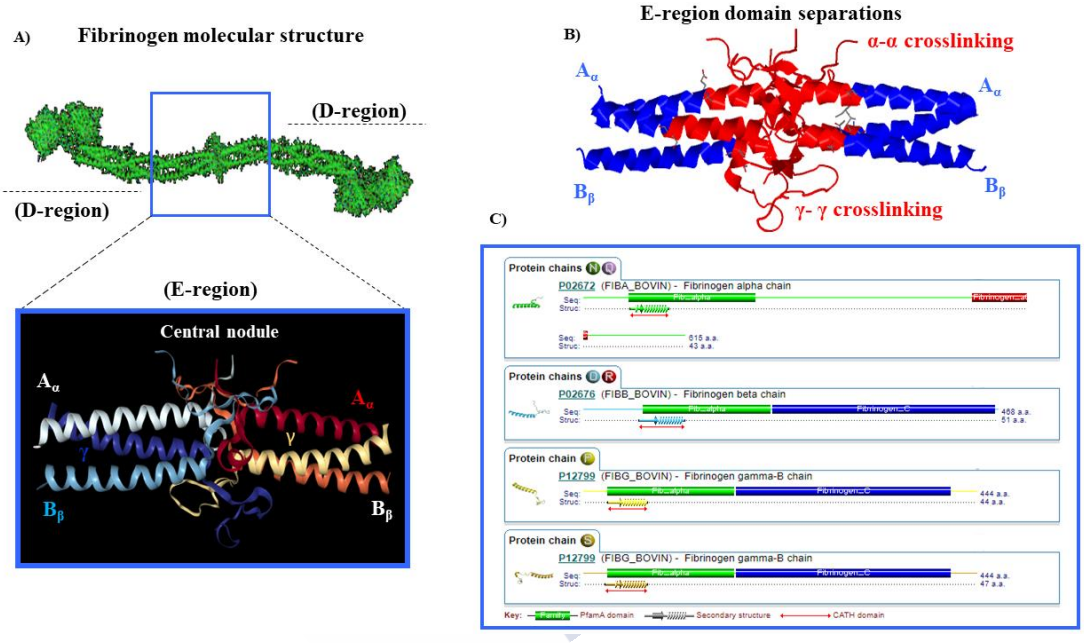


Figure 1. **A)** Representation of general structure of bovine fibrinogen by relevant regions like two C-terminal portions of the carbohydrate-linking coiled-coil (D-regions) and the central nodule (E-region) highlighted inside the blue rectangle with the corresponding chains (A_α and B_β chains) which make up a kind of funnel-hydrophobic cavity (thrombin binding-domain). **B)** Domain separation representation of E-region according to similarity structural and functional criteria showing the D-regions symmetric domains (color label blue) and the funnel-hydrophobic cavity formed by the γ-γ and α-α crosslinking from A_α and B_β chains (red-color label). **C)** Secondary structure and sequence of amino acid residue number by chains for the fibrinogen E-region. More details on the fibrinogen whole molecular structure can be found in **Figure S2.** Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

Next, to ensure success of molecular docking modeling one of the most important task is the prediction of feasible binding-sites of the protein, linked to an appropriate structural-validation and flexibility properties of the protein (fibrinogen E-region). Several algorithms and procedures for detecting relevant binding cavities have been implemented over the years to reach the cited goal. In this concern, in the present study the prediction of binding-sites from the bovine fibrinogen E-region was carried out by using machine learning methods-based 3D-Deep Convolutional Neural Networks. The results on the structure prediction/identification of fibrinogen binding site are presented in the **Figure 2**. Thus, it is possible to identify a single functional pocket in the protein structure. This pocket is surrounded by a remarkable abundance of polar residues as serine, threonine and cysteine ([CYS]48:Q.CA, [THR]21:P.CA, [THR]85:R.CA, [SER]50:Q.CB, [THR]85:Q.CA, [THR]22:S.CA, [CYS]44:N.C, [THR]85.O.CA). The distribution of these aminocids at the binding site is higher than in the rest of the protein structure. These residues play a primary role in catalysis and their presence restricts the possible types of interactions with ligands (propranolol and acebutolol). The flexibility of the protein is a crucial issue both for the correct prediction of the ligands binding modes and for the determination of the different conformational changes that the ligands can produce. This flexibility allows the opening, closing and adaptation of the binding sites to control the binding processes and the specific physiological function of the proteins. We have already commented that the residues surrounding the binding site determine their operation. However, residues outside the binding site can also have a long-range effect on the properties of the binding site[42-44]. The result of the calculation of the flexibility is shown in the **Figures 2** (**D** and **E**). As it can be observed the most rigid area corresponds to the binding site, while the rest of the structure is much more flexible. This fact enhances the importance of steric requirement's. In addition, the matrix cross-correlation shows how the distance between residues increases concomitantly to the reduction of their movements.

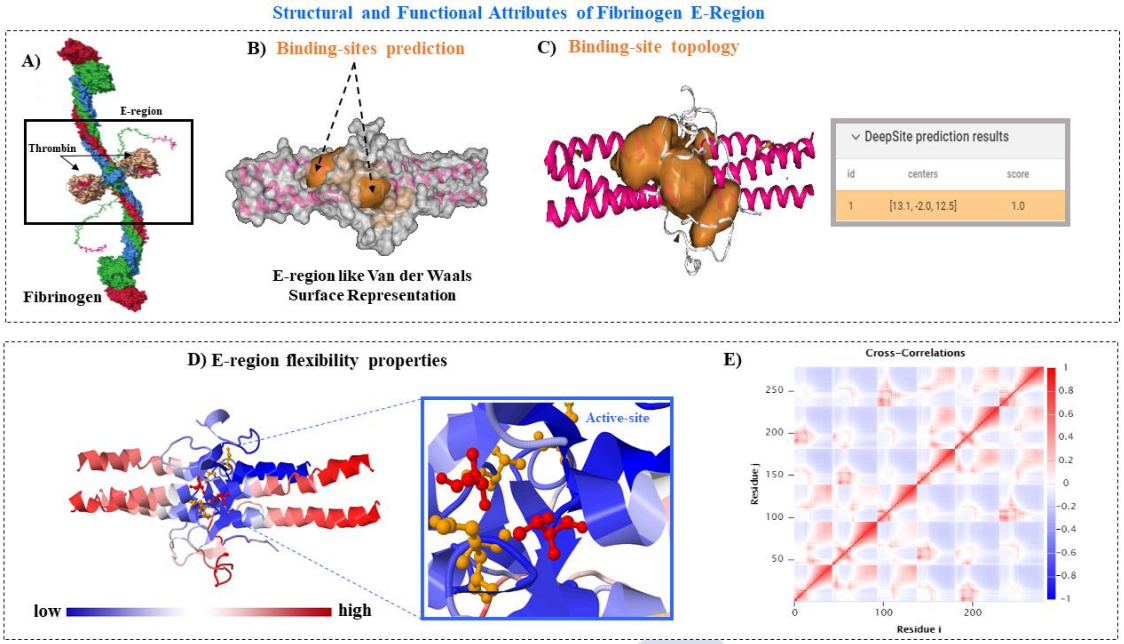


Figure 2. **A)** The top box of dashed lines is just to represent the structural attributes of bovine fibrinogen like: A) Representation of the full x-ray crystallographic molecular structure of bovine fibrinogen highlighting E-region with the corresponding thrombin binding-domain, **B)** and **C)** DeepSite convolutional neural network-based predictor for calculation of topological-cavities and dimensions of grid box center of docking simulation from fibrinogen E-region with maximum score = 1.0. The box down of dashed lines represent functional attributes of bovine fibrinogen (E-region) where C-terminal fragments (D-region) of the carbohydrate-linking coiled coils are omitted for clarity, then is depicted like: **D)** Flexibility analysis of fibrinogen E-region based on the fluctuations size of the slowest vibration-modes of the protein from E-region residues with low-flexibility (blue) to high-flexibility (red). Highlighting the E-region catalytic binding site (within the blue rectangle). **E)** Distance matrix-cross correlation showing the inter-residue flexibility-based correlations between residue fluctuations from E-region, plotted like a function of residue indices *i* and *j* (number of residue). Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

Following this sequence, to validate the 3D X-ray crystallographic quality of the E-region molecular structure model the Ramachandran plot was carried up as a control test to avoid false positive flexible docking experiments. The Ramachandran plot is a two-dimensional projection on the plane from 3D-crystallographic protein model that consider all the possible E-region conformations by residue; including the key binding sites residues

which are defined according to the torsion dihedral angles (*Psi vs. Phi*) around the fibrinogen peptide-bond residues. For this end, allowed torsion values for *Phi vs. Psi* dihedral angles are found inside the Ramachandran colored purple contour (*i.e.*, conformationally-favored E-region residues). Otherwise, are classified as E-region residues sterically-disallowed which are associated with torsion dihedral angles *Psi* vs. *Phi* values which appear outside of the Ramachandran colored purple contour (like conformationally non-favored E-region residues). See **Figure 3.** For this instance, was obtained a high-quality model structure based Ramachandran plot for fibrinogen E-region with high resolution of 1.4 Å, with expected % > 95 % (red dotted line) for the favored region of the residues.

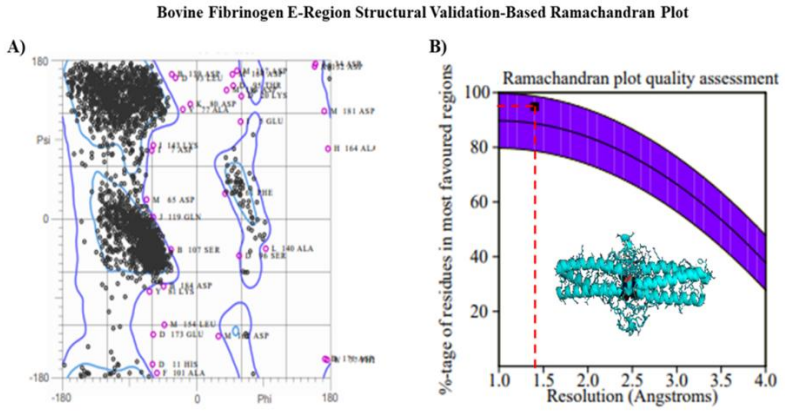


Figure 3. **A)** Representation of 2D-Ramachandran plot showing the Phi vs. Psi torsion dihedral angles and spatial distribution of Ramachandran outliers (empty pink circles) in the PDB x-ray whole structure of fibrinogen. Herein, all the possible torsion dihedral angles combinations of each amino-acid residue are depicted. **B)** Ramachandran plot quality assessment (fibrinogen E-region) is evaluated by the percentage (%) of the residues that are in the most-favored in the regions of the Ramachandran plot (or purple core) vs. crystallographic resolution (Å). To avoid obtaining false positives on the betablocker docking interactions, we check the absence of restricted-flexibility by filtering only the relevant betablocker interacting with binding residues from the Ramachandran plot of E-region catalytic binding site. Additional details on the Ramachandran plot just for E-region can be found like supplementary material (**Figure S3**) Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

Continuing the trend, we performed the molecular docking approaches to get the Gibbs free energy of binding (FEB, kcal/mol) for the docking complexes formed between the bovine fibrinogen E-region and the ligands (acebutolol and propranolol). For this instance, the docking results are classified as energetically-unfavorable when the FEB values ≥ 0 kcal/mol from the docking complexes, aiming at extremely low or complete absence of affinity of binding; otherwise is categorized as medium to high binding affinity. Then, modeling results show that the tested beta-adrenoreceptor blocking agents were able to interact-favorably with fibrinogen E-region following a spontaneous thermodynamic process (ΔG_{bind} < 0 kcal/mol) with very similar free energies as shown in the **Figure 4 (B** and **C)**. The local perturbation response scanning maps (LPRS maps) show the different patterns of interactions based on the strength of perturbation response with fibrinogen E-region binding residues in the absence (unoccupied fibrinogen E-region) and presence of ligands (acebutolol and propranolol). Highlighting the different behavior of local perturbation induced by the ligands in the 3D molecular structure of the protein in the top of the panel (inside the blue rectangle) with the corresponding intensity bar color (on the right). For this instance, from dark blue to red regions correspond to low to high strength of perturbation response respectively for 278 C(α)-atoms in the fibrinogen E-region. For this instance, all C(α)- atoms in the LPRS maps were established in range of the low frequency normal modes (from 1 to 3).

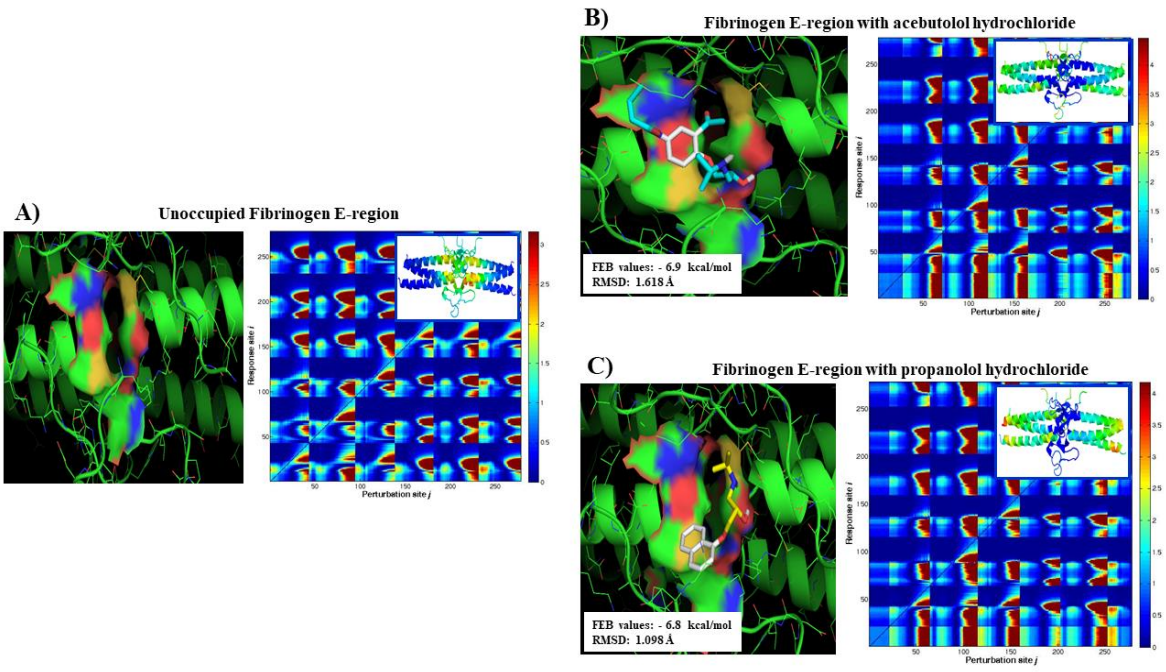


Figure 4. On the left of each panel van der Wall surface representation for the three cases studied like: **A)** unoccupied fibrinogen E-region, **B)** acebutolol/fibrinogen E-region docking complex, and **C)** propranolol/fibrinogen E-region docking complex with the corresponding Gibbs free energy of binding for the best crystallographic-docking poses (RMSD < 2Å). On the right of each panel is depicted the local perturbation response scanning maps (LPRS maps) for the aforementioned three cases. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

The elastic network approach from the best docking complexes was performed considering each fibrinogen E-region residues (*i.e.*, fibrinogen-node positions are identified with the coordinates of C(α)-atoms from N-residues) interacting with the ligands and the potential (U) is calculated like the sum of individual harmonic potentials[32-35]. In the present study the elastic network approach describes the potential *U* for ligand-fibrinogen E-region complex (*$fib_{(j)}$-$Lig_{(i)}$ or R-L*) as Hookean potential (or 'springs'):

$$U_{R-L} = \sum_{(i,j)\in N(R)\cup N(L)} \frac{1}{2}\gamma_R \left(d_{ij} - d_{ij}^0\right)^2 + \sum_{(i,j)\in N(,L)} \frac{1}{2}\gamma_L \left(d_{ij} - d_{ij}^0\right)^2 \qquad (4)$$

Herein, the force constant of the spring connecting ligand (*i*)-atoms and *j*-C(α)-atoms of the ligands (acebutolol and propranolol) and fibrinogen E-region receptor it is taken as γ_L and γ_R respectively. These values are defined as:

$$\gamma = e^{\left[\left(\frac{-1}{2}\right)\left(\left(\frac{d_{ij}^0}{d_{cutoff}}\right)^2\right)\right]} \qquad (5)$$

for this instance, the cutoff for γ_R = 1, d_{ij}^0 is the distance between *i*/*j*-C(α)-atoms of the ligand (*L*) and receptor (*R*) complexes, the equilibrium condition it is taken by the superscript "0" (assumed in elastic network model) and the d_{cutoff} = 7 Å. Next, for a system of N-atoms, there will be 3N-degrees of freedom, and $U_{R\text{-}L}$ is a function thereof. The 3N-degrees of freedom are taken from the Cartesian coordinates (*x*, *y*, and *z*) of each *i*, *j*-atom like atom charges: $(q_{ix}, q_{iy}, \text{and } q_{iz})_L$ and $(q_{jx}, q_{jy}, \text{and } q_{jz})_R$ for *L* and *R* respectively. In this concern normal mode analysis is based on harmonic potential approximation that allows to describe analytically solutions for vibrational motion-equations of unperturbed and/or perturbed protein under the influence of external stimuli (*i.e.*, beta-adrenoreceptor blocking ligand-interactions) through diagonalization of a Hessian matrix (U) like

potential energy matrix (U) obtained from mass-weighted second derivatives[38]. Where the normal modes are the eigenvectors of the Hessian matrix defined by:

$$U = \begin{bmatrix} U_{1,1} & U_{1,2} \cdots & U_{1,N} \\ U_{2,1} & U_{2,2} \cdots & U_{2,N} \\ \vdots & \vdots & \vdots \\ U_{N,1} & U_{N,2} \ldots & U_{N,N} \end{bmatrix} \quad (6)$$

where each U is a 3 (m + n) x 3 (m + n) matrix which holds the anisotropic information's regarding the orientation of residues (or number of C(α)-nodes of fibrinogen receptor (m or $R_{(m)}$) and ligands (n or $L_{(n)}$) respectively. Then, we can define U (U^C) of the potential ((U), *i*.e., $U_{R\text{-}L}$) according to:

$$U^C = \begin{bmatrix} U^R & U^{RL} \\ U^{LR} & U^L \end{bmatrix} \quad (7)$$

herein, U^C has two *3N x 3N* diagonal submatrices like U^R (3m x 3m) and U^L (3n x 3n) that describe the intramolecular coupling motions of ***R*** and ***L*** respectively[32-35]. These diagonal submatrices or super-elements is related to the coupled DoFs for the same atoms from the L_{ii} or R_{jj} for the evaluated ligands and fibrinogen respectively. The Hessian U^C diagonal super-elements (U^R *and* U^L) are calculated according to the following equation:

$$U_{i,i} = U^{RR} = U^{LL} = -\sum_{j=1, j\neq i}^{N} U_{i,j} \quad (8)$$

Furthermore, U^C has two *3N x 3N* off-diagonal submatrices like $U^{RL} \equiv (U^{LR})^T$ that describe the intermolecular docking interactions like perturbations responses between motions of *R* (*i.e.*, fibrinogen E-region *j*-C(α)-atoms) and ligands (*L:* acebutolol and propranolol)**.** For this instance, the Hessian matrix U^C for off-diagonal super elements (U^{RL} *and* $(U^{LR})^T$) present the form:

$$U_{i,j} = U^{RL} = \begin{bmatrix} \left(\frac{\partial^2 U_R}{\partial X_i \partial X_j} + \frac{\partial^2 U_L}{\partial X_i \partial X_j}\right) & \left(\frac{\partial^2 U_R}{\partial Y_i \partial X_j} + \frac{\partial^2 U_L}{\partial Y_i \partial X_j}\right) & \left(\frac{\partial^2 U_R}{\partial Z_i \partial X_j} + \frac{\partial^2 U_L}{\partial Z_i \partial X_j}\right) \\ \left(\frac{\partial^2 U_R}{\partial X_i \partial Y_j} + \frac{\partial^2 U_L}{\partial X_i \partial Y_j}\right) & \left(\frac{\partial^2 U_R}{\partial Y_i \partial Y_j} + \frac{\partial^2 U_L}{\partial Y_i \partial Y_j}\right) & \left(\frac{\partial^2 U_R}{\partial Z_i \partial Y_j} + \frac{\partial^2 U_L}{\partial Z_i \partial Y_j}\right) \\ \left(\frac{\partial^2 U_R}{\partial X_i \partial Z_j} + \frac{\partial^2 U_L}{\partial X_i \partial Z_j}\right) & \left(\frac{\partial^2 U_R}{\partial Y_i \partial Z_j} + \frac{\partial^2 U_L}{\partial Y_i \partial Z_j}\right) & \left(\frac{\partial^2 U_R}{\partial Z_i \partial Z_j} + \frac{\partial^2 U_L}{\partial Z_i \partial Z_j}\right) \end{bmatrix} \quad (9)$$

Herein, $i \in N$ (R) and $j \in \{k|\ (i, k) \in N\ (R, L)\}$ of the receptor (R) and ligands (L) respectively. The off-diagonal super-elements of can be replaced by the Hessian elements which represents the force of the perturbation (U^{RL}) induced by the ligands in a given magnitude on the fibrinogen E-region residues network as depicted below in equation 10:

$$U^{RL} = -\begin{bmatrix} \gamma_R\left(1+\frac{d_{ij}^0(x_j-x_i)^2}{(d_{ij})^3}-\frac{d_{ij}^0}{d_{ij}}\right)+\gamma_L\left(1+\frac{d_{ij}^0(x_j-x_i)^2}{(d_{ij})^3}-\frac{d_{ij}^0}{d_{ij}}\right) & \gamma_R\left(d_{ij}^0(y_j-y_i)\frac{x_j-x_i}{(d_{ij})^3}\right)+\gamma_L\left(d_{ij}^0(y_j-y_i)\frac{x_j-x_i}{(d_{ij})^3}\right) & \gamma_R\left(d_{ij}^0(z_j-z_i)\frac{x_j-x_i}{(d_{ij})^3}\right)+\gamma_L\left(d_{ij}^0(z_j-z_i)\frac{x_j-x_i}{(d_{ij})^3}\right) \\ \gamma_R\left(d_{ij}^0(x_j-x_i)\frac{y_j-y_i}{(d_{ij})^3}\right)+\gamma_L\left(d_{ij}^0(x_j-x_i)\frac{y_j-y_i}{(d_{ij})^3}\right) & \gamma_R\left(1+\frac{d_{ij}^0(y_j-y_i)^2}{(d_{ij})^3}-\frac{d_{ij}^0}{d_{ij}}\right)+\gamma_L\left(1+\frac{d_{ij}^0(y_j-y_i)^2}{(d_{ij})^3}-\frac{d_{ij}^0}{d_{ij}}\right) & \gamma_R\left(d_{ij}^0(z_j-z_i)\frac{y_j-y_i}{(d_{ij})^3}\right)+\gamma_L\left(d_{ij}^0(z_j-z_i)\frac{y_j-y_i}{(d_{ij})^3}\right) \\ \gamma_R\left(d_{ij}^0(x_j-x_i)\frac{z_j-z_i}{(d_{ij})^3}\right)+\gamma_L\left(d_{ij}^0(x_j-x_i)\frac{z_j-z_i}{(d_{ij})^3}\right) & \gamma_R\left(d_{ij}^0(y_j-y_i)\frac{z_j-z_i}{(d_{ij})^3}\right)+\gamma_L\left(d_{ij}^0(y_j-y_i)\frac{z_j-z_i}{(d_{ij})^3}\right) & \gamma_R\left(1+\frac{d_{ij}^0(z_j-z_i)^2}{(d_{ij})^3}-\frac{d_{ij}^0}{d_{ij}}\right)+\gamma_L\left(1+\frac{d_{ij}^0(z_j-z_i)^2}{(d_{ij})^3}-\frac{d_{ij}^0}{d_{ij}}\right) \end{bmatrix}$$

(10)

By the other hand, we define the (i,j)-residue fluctuation like ΔR_i and ΔR_j, we denote as R_i^0 and R_j^0 vectors respect to the *XYZ*-components of the instantaneous distance vector R_{ij}^0 for the equilibrium position and their corresponding instantaneous-values are R_i and R_j. Furthermore, d_{ij}^0 and d_{ij} represent the equilibrium position and instantaneous separation vectors between i and j residue sites. Separation changes (*i.e.*, displacements-based R-L perturbation) with respect to the equilibrium coordinates like $x_{ij}^0 = (x_j^0 - x_i^0)_{R-L}$, $y_{ij}^0 = (y_j^0 - y_i^0)_{R-L}$ and $z_{ij}^0 = (z_j^0 - z_i^0)_{R-L}$ are depicted by $\Delta d_{ij} = (d_{ij} - d_{ij}^0) = \Delta R_j - \Delta R_i$. For this instance, the cross-correlation between fibrinogen residue fluctuations can be obtained by means of the equation:

$$\langle \Delta R_i \times \Delta R_j \rangle = \frac{3k_B T}{\gamma}[U^{-1}]_{ij} \tag{11}$$

Where $U^{-1}{}_{ij}$ is the ij^{th} element of the inverse of the elastic network potential U^{RL} given by the equation (12). Therefore, the determination of cross-correlations between fibrinogen residue fluctuations is determined by the ij^{th} element of the U^{-1} of U [42-46]. Then, the normalized cross-correlation (C_{ij}) between the i^{th} and j^{th}-fibrinogen residue is defined by:

$$c_{ij} = \frac{\langle \Delta R_i \times \Delta R_j \rangle}{\left[\left\langle (\Delta R_i)^2 \times (\Delta R_j)^2 \right\rangle\right]^{\frac{1}{2}}} \tag{12}$$

In the LPRS maps the i^{th} rows is related to the response generated upon perturbing residue (i); its average (*i.e.*, over all receivers, k- effector residues of fibrinogen. Similarly, the k^{th}

columns of the LPRS map represent the sensitivity in response to the perturbation of all fibrinogen allosteric residues (*j*-sensors residues)[42-46].

Besides, the elastic network analysis on local perturbation of fibrinogen E-region residue pairs (*i, j*), allows to know the directions of movement, normal modes of fluctuations, of the atoms that make them up. These modes represent all the fluctuation patterns or possible ways in which the structure of the protein will oscillate around its equilibrium conformation (or minimum local energy) when perturbed by the present of ligands in the binding sites. The vibrational analysis of the proteins, allows to know the directions of movement, normal modes of vibration, of the atoms that make them up. These modes represent all the patterns or possible ways in which the structure of the protein will oscillate around its equilibrium conformation (minimum local energy) when disturbed. Deformation of a protein can be expressed as a linear combination of several of these calculated modes. The modes with higher frequencies represent very localized movements in which a low number of atoms intervene (vibration of the hydrogen atoms). On the other hand, the modes with the lowest frequencies correspond to movements delocalized or collective in which a large number of atoms are involved that oscillate with a considerable amplitude. Precisely these are the most relevant modes because they are related to the functional properties of proteins. In this regard, the fibrinogen deformation can be expressed as a linear combination of several of these calculated flexibility normal modes. According to this, the properties of elastic network models for the fibrinogen based on the frequency ($1/\lambda$) was determinated like the reciprocal of eigenvalue and collectivity degree of normal modes from the fibrinogen E-region which are affected by the presence of ligands inducing appreciable changes in the inter-residue communication (*i.e.*, correlation of the vibration by cluster de residues). According to the definition the collectivity degree (K_k) indicates the fraction of residues (fibrinogen E-region residues) that are significantly affected by a given *k*-normal mode. See **Figure 5**.

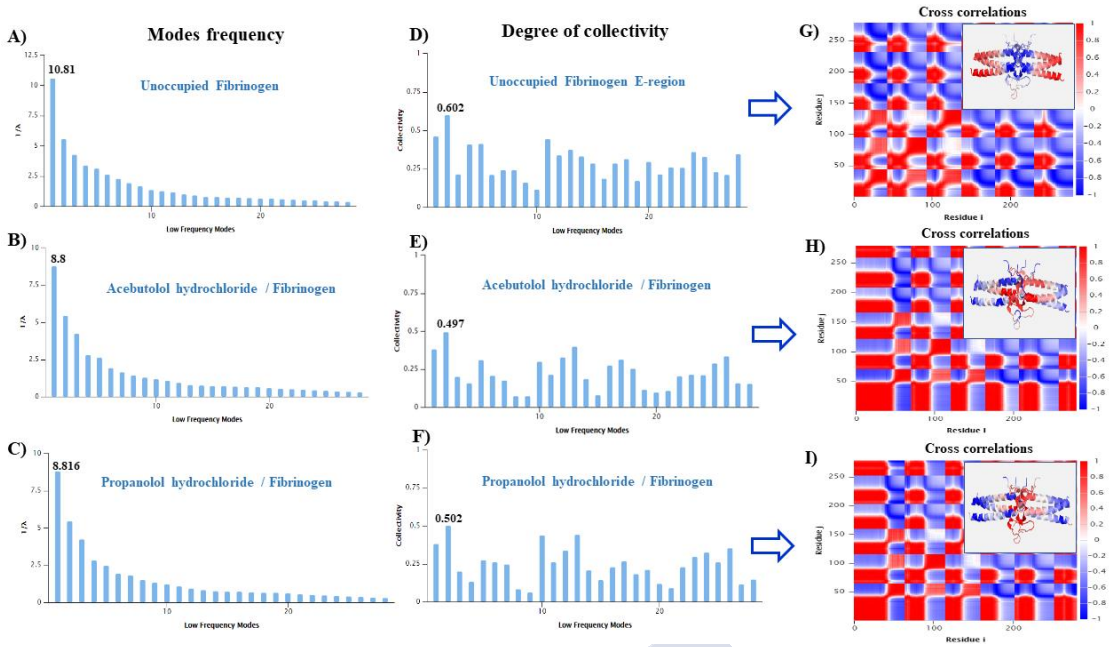


Figure 5. Graphical representation of the properties of the fibrinogen E-region elastic network models based on normal modes frequencies ($1/\lambda$) *vs.* low frequency normal modes for three evaluated like **A)** unoccupied fibrinogen E-region, **B)** acebutolol hydrochloride/fibrinogen and **C)** propranolol hydrochloride/fibrinogen. Next is represented the biophysical parameter degree of collectivity (K_k) *vs.* low frequency normal modes for three evaluated conditions: **D)** unoccupied fibrinogen E-region, **E)** and **F)** for acetabutolol and propranolol respectively. On the far right are represented the matrix-cross correlation maps showing the correlations between residue fluctuations from fibrinogen E-region, which are plotted like a function of *i vs. j* residue indices together with the corresponding flexibility changes in the 3D molecular structure of fibrinogen E-region (in the top rectangle) in the absence and presence of beta-adrenoreceptor blocking ligands like **G)** unoccupied fibrinogen E-region, **H)** acebutolol hydrochloride/fibrinogen and **I)** propranolol hydrochloride/fibrinogen. The *i,j*-pairs to fully correlated fluctuation-residues with strong correlations ($0.2 \leq C_{ij} \leq 1$) with color red-label pointing the same orientation for the *i* and *j* fibrinogen residues-flexibility. while the anti-correlated motion-based flexibility (*i.e.*: $C_{ij} \leq 0$) are colored in blue (*i.e.*, opposite-direction for the *i,j*-residue fluctuations and for moderately correlated and/or uncorrelated fibrinogen regions ($C_{ij} \approx 0$) are colored in light red and blue, respectively. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

The motion frequency of normal modes for unoccupied fibrinogen E-region and in the presence of beta-adrenoreceptor blocking ligands have been evaluated by the reciprocal

of the eigenvalue (1/λ)[42-46]. From the biophysical point of view, a higher value of 1/λ indicates a lower mode with low motion frequency which are highly associated to the biochemical functions and represent very localized movements in which a low number of atoms from *R* and *L* are involved. The present of the ligands acebutolol and propranolol in the fibrinogen E-region binding site showed a decrease following the reference of the maximum frequency eigenvalue (1/λ: from 10.81 to 8.8 and 8.816) suggesting at least theoretically a loss of fibrinogen function for both cases. The same behavior was observed for the biophysical parameter collectivity degree where was obtained a value of unoccupied fibrinogen E-region (0.602) > propranolol (0.502) > acebutolol (0.497) following the reference of maximum degree of collectivity. This parameter informs about how the structural elements (fibrinogen binding residues) move together in that particular mode. The collectivity (K_k) for a given *k*-normal mode is estimate of the cooperativity degree between residues in that normal mode and can be defined according to:

$$K_{K(R-L)} = \frac{1}{N} e^{-\sum_i^N (\Delta d_{ij})^2 ln(\Delta d_{ij})^2} \quad (13)$$

where, *k* is the normal mode number, $\Delta d_{ij} = d_{ij} - d_{ij}^0$ is the magnitude of the displacement (*i.e.*, displacements-based R-L perturbation) previously defined tacking into account the C(α)-atoms from residue in the mode *k*. The subscripts *i* and *j* are the indices for interacting *i*,*j*-fibrinogen residues.

Then, a high collectivity degree is associated to high entropy, that means highly-cooperative normal mode that involving a significant region or sometimes the full structure of the protein (*i.e.*, ligand interaction inducing motion-fluctuations distributed over a larger number of fibrinogen residues rather than being orderly confined to a few residues). From the biophysical point of view lower frequency (1/λ) usually present higher collectivity degree that triggering cooperative allosteric-responses which are mainly of interest to explain docking interactions based on perturbation induced by ligand in allosteric binding residues of fibrinogen molecule[42-46].

According to the **Figure 5 (G**, **H**, and **I)** the theoretical evidences suggest that the molecular docking mechanism of the studied beta-adrenoreceptor blocking ligands with the fibrinogen molecule are most likely based on local-perturbations of normal modes of vibration of allosteric residues that are part of the funnel-hydrophobic cavity of E-region. Particularly affecting the flexibility properties of the fibrinogen central nodule formed by

the γ-γ and α-α crosslinking from A_α and B_β chains by increasing the dynamical correlation between the residue fluctuations which interact with the acebutolol and propranolol. These effects theoretically could have negative consequences for the coagulation process (fibrin polymerization) preventing the assembly of the thrombin molecule. The individualized correlation analysis by elastic normal modes has a paramount importance in the identification of long-range perturbation-based correlations. In fact, in the coordinate space, the spatial correlation between the positions of *i*, *j*-residue pairs contain the summation of all modes (inter-modes cooperation) for a given cutoff motion frequency. That is meanwhile, the extent of inter-modes cooperation decays notably with increasing displacements-based R-L perturbation (Δd_{ij}) and by the presence of beta-adrenoreceptor blocking ligands that could affect the functional states of fibrinogen molecule which are strongly dependent on its flexibility properties.

In order to get more fully acquainted with the frequency-collectivity degree relationship ($1/\lambda \leftrightarrow K_{(k)}$) a clustering approach was performed by considering the aforementioned three conditions (Unoccupied fibrinogen E-region, acebutolol hydrochloride/fibrinogen and propranolol hydrochloride/fibrinogen). See **Figure 6.**

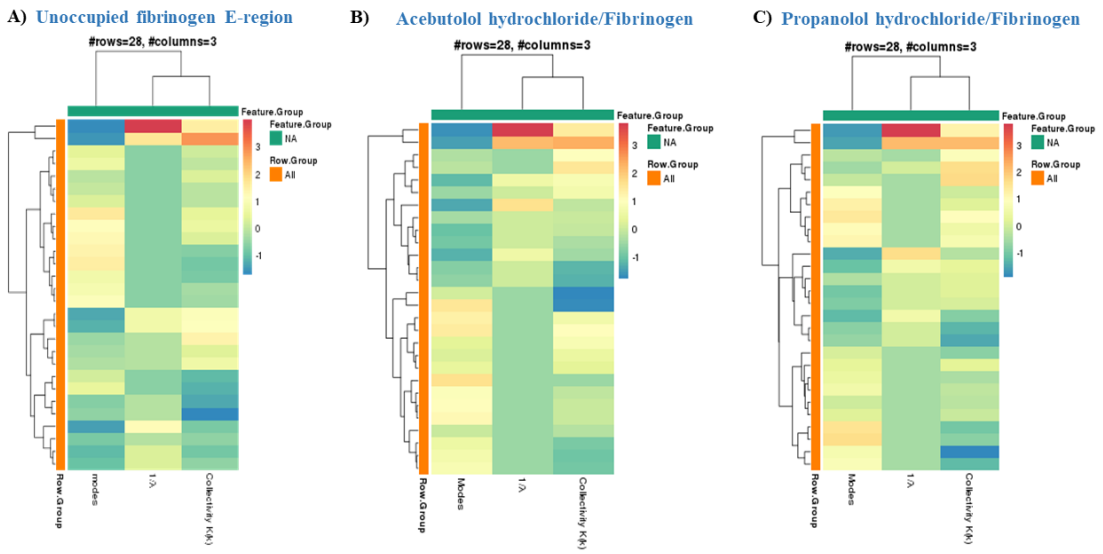


Figure 6. Hierarchical cluster analysis (HCA) based on K-means clustering for frequency-collectivity degree relationship ($1/\lambda \leftrightarrow K_{(k)}$) considering evaluated modeling conditions like **A)** unoccupied fibrinogen E-region, **B)** acebutolol hydrochloride/fibrinogen and **C)** propranolol hydrochloride/fibrinogen. The color intensity bar on the right represent the cluster dissimilarity based on the correlation

coefficients of the aforementioned conditions. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

In this context, it is possible to extract detailed information about hidden large-scale of relationships that affect the fibrinogen binding interaction with the beta-adrenoreceptor blocking ligands by considering the properties of normal modes and extracting similarities and differences on the interaction patterns to explain the binding mechanisms. To this end, we applied hierarchical cluster analysis based on K-means clustering with heatmaps following the criteria of maximum average dissimilarity. [47-50] This approach allows to visualize high-dimensional data because aims to partition *n*-observations into *k*-clusters where each observation is part of the cluster with the nearest mean, which serves as a prototype of it. Identifying *k*-number of centroids, and after that assigns each data point to the nearest cluster referred to averaging of the *n*-observations (i.e., finding the centroid)[47-50]. The *k*-means algorithm is an unsupervised machine learning algorithm that can be applied for this instance to efficient partition of the input data set formed by (# of normal modes, $1/\lambda$ and $K_{(k)}$)) into *k*-clusters including the values of data points and the correlation coefficients between the *n*-observations (*i.e.*, rows 28: categories or clusters for the $1/\lambda \leftrightarrow K_{(k)}$ relationships by individual layers) and the evaluated factors like individual normal mode properties (*i.e.*, 3 columns: # of normal modes, frequency and degree of collectivity) are simultaneously represented in the same graphs[47-50]. Herein, although the evaluated drugs (acebutolol and propranolol) belong to the same pharmacological group (beta-adrenoreceptor blocking agents) it is easy to identify cluster dissimilarity in the interaction-patterns from the physiological condition represented by unoccupied fibrinogen E-region. Particularly, the elastic network parameters $1/\lambda$ and $K_{(k)}$ which were presented in the same cluster based on y-axis (columns of factors or feature group) showing marked cluster-dissimilarity based on the correlation coefficients of the evaluated conditions with beta-adrenoreceptor blocking agents acebutolol and propranolol and the unoccupied fibrinogen E-region and also between these two drugs when we analyze the number and of clusters and hierarchical architecture individually represented by the x-axis to categorize the responses or relationship evaluated[47-50].

References

[1] R. Harada, N. Tochio, T. Kigawa, Y. Sugita, M. Feig, Reduced native state stability in crowded cellular environment due to protein–protein interactions, Journal of the American Chemical Society, 135 (2013) 3696-3701.

[2] A. Gonzalez-Perez, J. Ruso, G. Prieto, F. Sarmiento, Physicochemical study of ovalbumin in the presence of sodium dodecyl sulphate in aqueous media, Colloid and Polymer Science, 282 (2004) 351-356.

[3] J.M. Ruso, N. Deo, P. Somasundaran, Complexation between dodecyl sulfate surfactant and zein protein in solution, Langmuir, 20 (2004) 8988-8991.

[4] A. Bratek-Skicki, P. Żeliszewska, J.M. Ruso, Fibrinogen: a journey into biotechnology, Soft Matter, 12 (2016) 8639-8653.

[5] N. Hassan, L.R. Barbosa, R. Itri, J.M. Ruso, Fibrinogen stability under surfactant interaction, Journal of colloid and interface science, 362 (2011) 118-126.

[6] S. Amin, Mechanical factors and bone health: effects of weightlessness and neurologic injury, Current rheumatology reports, 12 (2010) 170-176.

[7] L. Feng, J.D. Andrade, Structure and Adsorption Properties of Fibrinogen, in: Proteins at Interfaces II, vol. 602, American Chemical Society, 1995, pp. 66-79.

[8] J. de Azevedo, F. Walter, R.A. Caceres, I. Pauli, L.F.S. Timmers, G.B. Barcellos, K.B. Rocha, M.B. Soares, Protein-drug interaction studies for development of drugs against Plasmodium falciparum, Current drug targets, 10 (2009) 271-278.

[9] J.M. Ruso, P. Taboada, P. Martínez-Landeira, G. Prieto, F. Sarmiento, A comparative study of the interaction between nafcillin and catalase by equilibrium dialysis and ζ-potential measurements, The Journal of Physical Chemistry B, 105 (2001) 2644-2648.

[10] N. Hassan, J.M. Ruso, P. Somasundaran, Mechanisms of fibrinogen–acebutolol interactions: Insights from DSC, CD and LS, Colloids and Surfaces B: Biointerfaces, 82 (2011) 581-587.

[11] N. Hassan, J. Maldonado-Valderrama, A.P. Gunning, V. Morris, J.M. Ruso, Investigating the effect of an arterial hypertension drug on the structural properties of plasma protein, Colloids and Surfaces B: Biointerfaces, 87 (2011) 489-497.

[12] E.C. Meng, B.K. Shoichet, I.D. Kuntz, Automated docking with grid-based energy evaluation, Journal of Computational Chemistry, 13 (1992) 505-524.

[13] M. Feig, A. Onufriev, M.S. Lee, W. Im, D.A. Case, C.L. Brooks III, Performance comparison of generalized born and Poisson methods in the calculation of electrostatic

solvation energies for protein structures, Journal of Computational Chemistry, 25 (2004) 265-284.

[14] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic acids research, 28 (2000) 235-242.

[15] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, Journal of computational chemistry, 31 (2010) 455-461.

[16] S. Forli, R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, A.J. Olson, Computational protein–ligand docking and virtual drug screening with the AutoDock suite, Nature protocols, 11 (2016) 905.

[17] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, PubChem substance and compound databases, Nucleic acids research, 44 (2015) D1202-D1213.

[18] J. de Ruyck, G. Brysbaert, R. Blossey, M.F. Lensink, Molecular docking as a popular tool in drug design, an in silico travel, Advances and applications in bioinformatics and chemistry: AABC, 9 (2016) 1.

[19] K.M. Elokely, R.J. Doerksen, Docking challenge: protein sampling and molecular docking performance, Journal of chemical information and modeling, 53 (2013) 1934-1945.

[20] G. Heberlé, W. F de Azevedo, Bio-inspired algorithms applied to molecular docking simulations, Current medicinal chemistry, 18 (2011) 1339-1352.

[21] G. Heberlé, W. F de Azevedo, Bio-inspired algorithms applied to molecular docking simulations, Current medicinal chemistry, 18 (2011) 1339-1352.

[22] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallographica Section D: Biological Crystallography, 66 (2010) 12-21.

[23] J. Jiménez, S. Doerr, G. Martínez-Rosell, A. Rose, G. De Fabritiis, DeepSite: protein-binding site predictor using 3D-convolutional neural networks, Bioinformatics, 33 (2017) 3036-3042.

[24] W.P. Feinstein, M. Brylinski. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets, Journal of cheminformatics, 7 (2015) 18.

[25] A. Tao, Y. Huang, Y. Shinohara, M.L. Caylor, S. Pashikanti, D. Xu, ezCADD: A Rapid 2D/3D Visualization-Enabled Web Modeling Environment for Democratizing Computer-Aided Drug Design, Journal of chemical information and modeling, 59 (2018) 18-24.

[26] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, in, ACS Publications, 2011.

[27] R. Thomsen, M.H. Christensen, MolDock: a new technique for high-accuracy molecular docking, Journal of medicinal chemistry, 49 (2006) 3315-3321.

[28] B. Kramer, M. Rarey, T. Lengauer, CASP2 experiences with docking flexible ligands using FlexX, Proteins: Structure, Function, and Bioinformatics, 29 (1997) 221-225.

[29] J. Liu, R. Wang, Classification of current scoring functions, Journal of chemical information and modeling, 55 (2015) 475-482.

[30] S. Mitternacht, I.N. Berezovsky, Coherent conformational degrees of freedom as a structural basis for allosteric communication, PLoS computational biology, 7 (2011) e1002301.

[31] O. Keskin, S.R. Durell, I. Bahar, R.L. Jernigan, D.G. Covell, Relating molecular flexibility to function: a case study of tubulin, Biophysical journal, 83 (2002) 663-680.

[32] J.G. Greener, M.J. Sternberg, AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis, BMC bioinformatics, 16 (2015) 335.

[33] W. Zheng, S. Doniach. A comparative study of motor-protein motions by using a simple elastic-network model, Proceedings of the National Academy of Sciences, 100 (2003) 13253-13258.

[34] U. Emekli, D. Schneidman-Duhovny, H.J. Wolfson, R. Nussinov, T. Haliloglu, HingeProt: automated prediction of hinges in protein structures, Proteins: Structure, Function, and Bioinformatics, 70 (2008) 1219-1227.

[35] T. Oliwa, Y. Shen, cNMA: a framework of encounter complex-based normal mode analysis to model conformational changes in protein interactions, Bioinformatics, 31 (2015) i151-i160.

[36] J. Madrazo, J.H. Brown, S. Litvinovich, R. Dominguez, S. Yakovlev, L. Medved, C. Cohen, Crystal structure of the central region of bovine fibrinogen (E5 fragment) at 1.4-Å resolution, Proceedings of the National Academy of Sciences, 98 (2001) 11967-11972.

[37] I. Pechik, S. Yakovlev, M.W. Mosesson, G.L. Gilliland, L. Medved, Structural basis for sequential cleavage of fibrinopeptides upon fibrin assembly, Biochemistry, 45 (2006) 3588-3597.
[38] J.W. Weisel, G.N. Phillips Jr., C. Cohen, THE STRUCTURE OF FIBRINOGEN AND FIBRIN: II. ARCHITECTURE OF THE FIBRIN CLOT*, Annals of the New York Academy of Sciences, 408 (1983) 367-379.
[39] J.W. Weisel, G.N. Phillips, C. Cohen, The structure of fibrinogen and fibrin: II. Architecture of the fibrin clot, Annals of the New York Academy of Sciences, 408 (1983) 367-379.
[40] R. Hantgan, J. McDonagh, J. Hermans, FIBRIN ASSEMBLY*, Annals of the New York Academy of Sciences, 408 (1983) 344-366.
[41] J.H. Brown, N. Volkmann, G. Jun, A.H. Henschen-Edman, C. Cohen, The crystal structure of modified bovine fibrinogen, Proceedings of the National Academy of Sciences, 97 (2000) 85-90.
[42] A. Stank, D.B. Kokh, J.C. Fuller, R.C. Wade, Protein Binding Pocket Dynamics, Accounts of Chemical Research, 49 (2016) 809-815.
[43] A.R. Atilgan, S. Durell, R.L. Jernigan, M. Demirel, O. Keskin, I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, Biophysical journal, 80 (2001) 505-515.
[44] I. Bahar, T.R. Lezon, A. Bakan, I.H. Shrivastava, Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins, Chemical reviews, 110 (2009) 1463-1497.
[45] C. Atilgan, A.R. Atilgan, Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein, PLoS computational biology, 5 (2009) e1000544.
[46] D. Penkler, O.z. Sensoy, C. Atilgan, O.z. Tastan Bishop, Perturbation-Response Scanning Reveals Key Residues for Allosteric Control in Hsp70, Journal of chemical information and modeling, 57 (2017) 1359-1374.
[47] A. Fernández, S. Gómez. Solving non-uniqueness in agglomerative hierarchical clustering using multidendrograms, Journal of Classification, 25 (2008) 43-65.
[48] G.J. Szekely, M.L. Rizzo, Hierarchical clustering via joint between-within distances: Extending Ward's minimum variance method, Journal of classification, 22 (2005) 151-183.
[49] O. Maimon, L. Rokach, Data mining and knowledge discovery handbook, (2005).

[50] D. Chen, L.-Y. Fu, D. Hu, C. Klukas, M. Chen, K. Kaufmann, The HTPmod Shiny application enables modeling and visualization of large-scale biological data, Communications biology, 1 (2018) 89.

Supplementary material

Molecular structure of beta-blockers drugs:

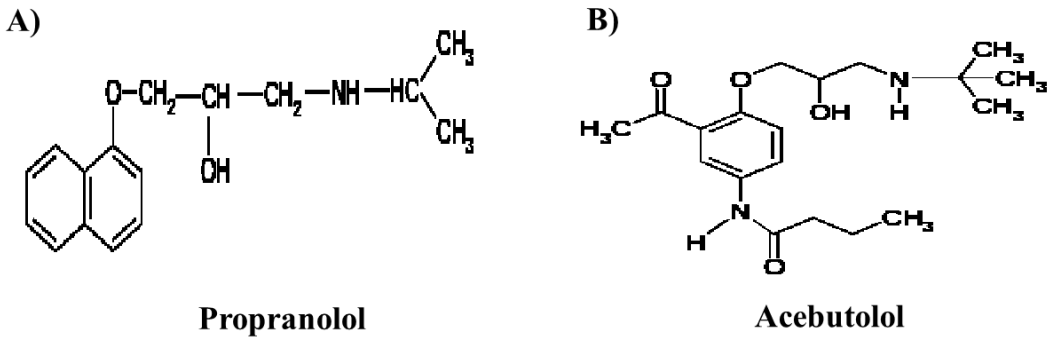


Figure S1. Molecular structure of the betablocker drug under study as **A)** Propranolol and **B)** Acebutolol. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

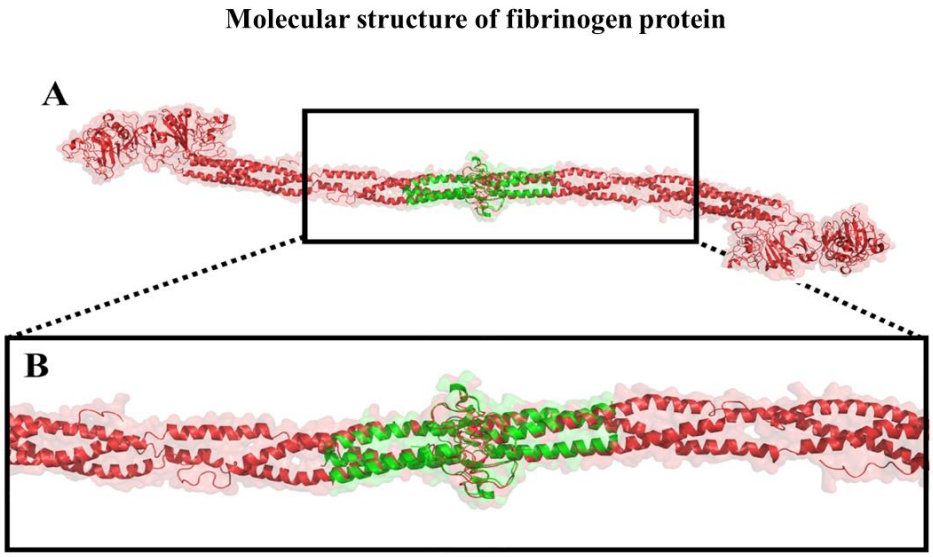


Figure S2. Molecular structure of the fibrinogen protein as **A)** whole molecular structure (marked like red plus green color), and **B)** Fibrinogen structure highlighting the E-region marked in green color. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

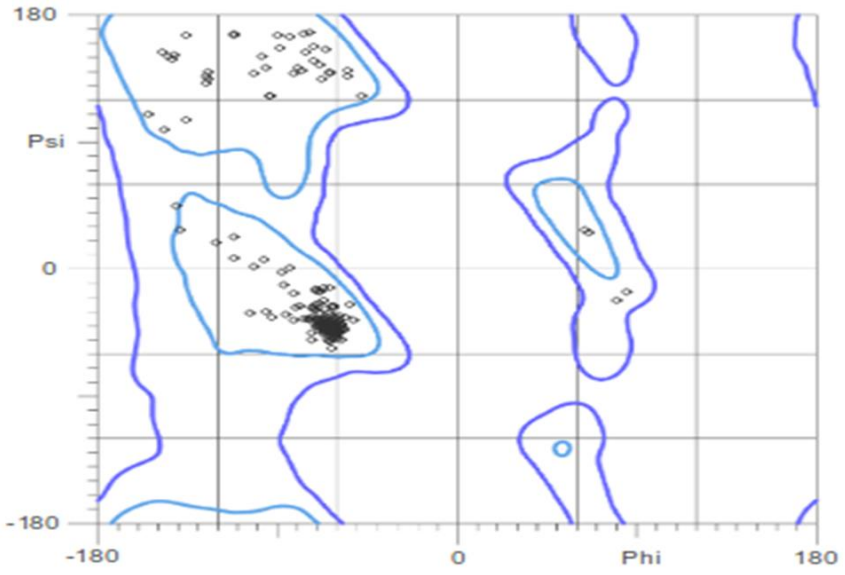


Figure S3. Representation of 2D-Ramachandran plot showing the Phi vs. Psi torsion dihedral angles for the molecular structure of fibrinogen E-region binding site. Herein, all the possible torsion dihedral angles combinations of each amino-acid residue are

depicted as black dots within the purple contour line. For this instance, note the total absence of Ramachandran outliers for conformationally-unfavored amino acid residues (i.e., outsite the purple contour line). Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2019.06.229 (accessed on 15 september 2019)

Chapter 4.

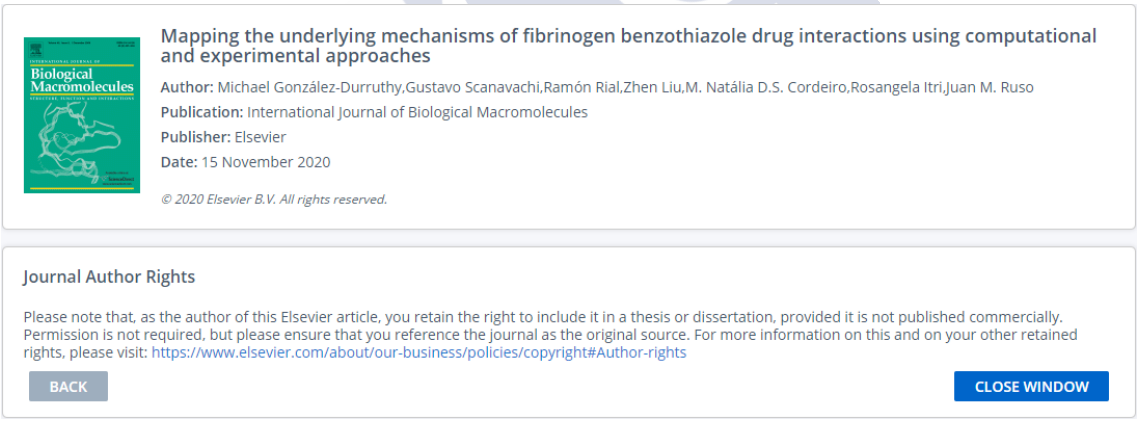
Michael González-Durruthy, Gustavo Scanavachi, Ramón Rial, Zhen Liu, M. Natália D. S. Cordeiro, Rosangela Itri, Juan M. Ruso. Mapping the underlying mechanisms of fibrinogen benzothiazole drug interactions using computational and experimentais approaches, *International Journal of Biological Macromolecules*, 163, 730-744, 2020. doi.org/10.1016/j.ijbiomac.2020.07.044.

Journal Impact Factor (JIF) (2020): 6.953.

CiteScore (2020): 8.5

Author Contribution M.G.-D.; Conceptualization, methodology-based on molecular docking, Fourier transform based on BTS-ligand conformational analysis, perturbation models, fractal approaches, and writing—original draft preparation.

Journal authorization:



Chapter 4.

Mapping the underlying mechanisms of fibrinogen benzothiazole drug interactions using computational and experimental approaches

Abstract

Three-dimensional conformational crystallographic binding-modes are of paramount importance to understand the docking mechanism of protein-ligand interactions and to identify potential "*leading drugs*" conformers towards rational drugs-design. Herein, we present an integrated computational-experimental study tackling the problem of multiple binding modes among the ligand 3-(2-Benzothiazolylthio)-propane sulfonic acid (BTS) and the fibrinogen receptor (E-region). Based on molecular docking simulations, we found that the free energy of binding values for nine of different BTS-docking complexes (*i.e.,* BTS-*pose_1-9*) were very close. We have also identified a docking-mechanism of BTS-interaction mainly based on non-covalent hydrophobic interactions with H-bond contacts stabilizing the fibrinogen-BTS docking complexes. Interestingly, the different BTS-*poses_1-9* were found to be able to block the fibrinogen binding site (E-region) by inducing local perturbations in effector and allosteric residues, reducing the degree of collectivity in its flexibility normal modes. As such, we theoretically suggest that the BTS-binding modes can significantly affect the physiological condition theoretically represented by the unoccupied fibrinogen protein structure by bringing global and local perturbations in the frequency domain spectra. The theoretical conformational mechanisms of interactions proposed here, were further corroborated by applying isothermal titration calorimetry (ITC) and small angle X-ray scattering (SAXS) experiments. Overall theoretical and experimental results shall open new avenues towards the application of complex supra-molecular information toward rational drugs-design.

4.1 Introduction

Proteins are responsible for most of the functions that occur in living organism, and for that purpose they need to maintain their three-dimensional structure. However, this is not an easy task because they are thermodynamically active, highly dynamic and tend to interact with several competing targets [1-3]. In this context, the arrangement by which proteins interact

with other molecular entities, whether specifically or not, has challenged many researchers to define biophysical parameters that characterize this action and has generated an authentic avalanche of publications [4, 5]. Nevertheless, still nowadays, support in this topic is in great demand as a result of increased awareness in biotechnology and nanomedicine applications. Despite all the advances in this field, establishing a protein's structure in its native state is a complex task and many physical methods are cumulatively required for its elucidation. In the last decades, computational tools have been added to complement the toolkit of scientists. However, it is not always easy to work with these two different tandem approaches, but a balance is necessary. In order for the approaches to remain broad and varied, researchers must not allow any of them to be eternally petrified. As always, the work of the researchers is to shed light to the enormous unknowns that hinder the path to the total understanding of the phenomena. Always bearing in mind that the results obtained by themselves must be sufficiently solid to be recognized while flexible enough in order not to limit future knowledge [6, 7].

Many diseases are based on biomolecular interactions. The different interactions of the proteins trigger mechanisms and signals that can indicate the state of health or disease in the organisms. Thus, protein interactions are behind the molecular basis of diseases and might assist methods of prevention, diagnosis and treatment [8]. On the other hand, the increase in the clinical use of therapeutic proteins has drawn attention to the potential for protein-drug interactions [9] and regulatory agencies have recently issued interaction guidelines, which require determination of plasma protein binding [10].

Fibrinogen (or factor I) is a large and elongated triglobular protein, part of the blood plasma of vertebrates. Its main function is to form fibrin clots to avoid the loss of blood due to injury. The protein (around 340 kDa) is 45 nm long, with a diameter of 5 nm, comprising a round disulfide-linked dimer composed of three pairs of nonidentical polypeptide chains, called Aα, Bβ, and γ (attached with 29 disulfide bonds and no free sulfhydryl groups), the predominant forms of which have 610, 461, and 411 amino acids, respectively [11]. Plasma levels of fibrinogen have been described as risk predictors of cardiovascular disease. It has been shown that drugs such as ticlopidione or fibrates are effective reducing fribrinogen levels. In addition, this effectiveness is affected by various factors, such as the presence of other diseases like diabetes [12]. Curcumin is a popular pigment with numerous known medical applications. However, its rapid degradation in biological fluids limits its action. Leung *et al.* have shown that the binging of curcumin to fibrinogen suppresses the degradation process by 95% [13]. Further, nanoparticles have been demonstrated to be very important for drug delivery systems

and, consequently, highlighting the role of protein-nanoparticle interactions [14]. In that respect, the binding of fibrinogen with gold nanoparticles has also been reported, underlining the importance of the process dynamics [15]. Computational techniques have also been focused on fibrinogen. Liu *et al*. have reported molecular docking simulations of the interaction of an anticoaagulant drug and fibrinogen to obtain features of the possible binding models. They found that hydrogen bonding and hydrophobic interactions embody 86% of the total energy [16].

Benzothiazole derivatives have been extensively applied in both therapeutic treatments and diagnosis including amyloid-β markers and dyes to follow protein conformational changes in cells. It has been discovered that a series of benzothiazoles substitutes are selective and high affinity inhibitors of different triosephosphate isomerases (TIM) including Tripanosome cruzi (Tc) and Tripanosome brucei (Tb) TIMs. Among the anti-tumor drugs discovered recently, several studies have concluded that the benzothiazole nucleus has a strong anticancer activity against human cancers. Some authors have found that some compounds including the benzothiazol nucleus killed cells in a tumor-specific way. Benzothiazoles derivatives have indeed been reported to possess potent anticancer properties due to their structural similarity with naturally occurring purines as they can easily interact with biomolecules of the living systems [17]. Diverse biological activities such as antiviral, antifungal or antibacterial have also been described [18, 19].

In this light, this work aims at disclosing the interactions between fibrinogen and 3-(2-Benzothiazolylthio)-propanesulfonic acid (BTS). In order to establish different approaches and strategies to deal with this system, emphasis has been placed on innovative approaches. Thus, molecular docking and local perturbation response scanning maps were chosen, since these give information about the interaction energy and local protein conformational changes [20, 21]. Further, we introduce for the first time, a combined methodology by using two recognized image-processing algorithms to tackle the ligand-conformations problem of BTS from a qualitative and quantitative point of view, namely: Fast Fourier transform (2D-FFT) and fractal approaches [22, 23]. These two methods offered information about the supramolecular complex. On the other hand, experimental results provided by isothermal titration calorimetry and Small Angle X-ray scattering were analyzed and contrasted with the computational ones.

4.2 Material and Methods

4.2.1 Computational approaches

4.2.2 Computational modeling-based molecular docking. To study the binding properties between BTS and fibrinogen protein (E-region) a theoretical mechanistic study based on molecular docking was performed. Towards such end, first we prepared the protein receptor file (*i.e.*, fibrinogen E-region), which was withdrawan from the *RCSB Protein Data Bank* (PDB) X-ray structures [24], *i.e.* with *PDB ID*: 1JY2. Afterwards, the fibrinogen (E-region) receptor molecular structure was optimized by applying the AutoDock Tools 4 software, using the AutoDock Vina scoring function [25,26]. This software starts by removing all the crystallographic waters as well as all co-crystallized ligands if present. Next, H-atoms were included in the fibrinogen protein structure, following an appropriate hybridization geometry by adding Gasteiger-Marsili empirical atomic partial-charges and protonation states of the X-ray fibrinogen pdb structure [24]. Following this, the BTS ligand was obtained from the *Pubchem Data Base Chemical Structure Search* (PubChem CID: 162569; MF: $C_{10}H_{11}NO_3S_3$). [27]. The drug geometry optimization was carried out by using the MOPAC extension based on NDDO aproximation [25,26]. To study the drug docking mechanisms with bovine fibrinogen E-region, we used the Autodock Vina scoring function developed by Trott *et al*. [25,26] to obtain the free energy of binging (FEB) based on X-Score function which approximates the standard chemical potentials. Before the docking modeling, the fibrinogen E-region binding-sites were predicted through the ezPocket with *fconv* for binding-site detection [28-29]. This step was carried out as delimiting the access to fibrinogen-cavities, like van der Waals surfaces that are likely to bind to the ligands (BTS). To this end, the *fconv* analysis uses Delaunay triangulation with weighted points to detect plausible binding pockets. Herein, the volumetric map of the fibrinogen binding-site is generated together with the Cartesian *XYZ*-coordinates for the docking box simulations [28-31], *i.e.*: grid box size with dimensions of X= 64Å, Y= 52Å, Z= 52Å and grid box center X= 13.1Å, Y= -2.0Å, Z= 12.5Å. For this purpose, a docking accuracy was fixed at 100 and after that, the best nine conformations or BTS binding-poses were selected [28-31]. The docking affinity (FEB values) were classified like energetically-unfavorable when the FEB of fibrinogen-BTS complexes ≥ 0 kcal/mol, therefore indicating either extremely low or complete absence of affinity; otherwise the fibrinogen-BTS docking complexes were classified like medium to high docking affinity [25,26].

4.2.3 3D Lig-Plot diagrams. This algorithm was applied to evaluate the influence in the FEB values for the different contributions of the ligand binding-poses (BTS binding-poses = 9) interacting with the fibrinogen E-region. To specified the key relevant intermolecular interactions between BTS binding-poses with the fibrinogen, 3D Lig-Plot diagrams were analyzed. To tackle this objective, a software named ezLigPlot was used. This software determines the non-covalent intermolecular interactions present in a given protein-ligand complex and it builds automatically a 3D-interaction diagram that includes hydrophobic, H-bond, cation-Π, and Π-Π stacking interactions along with its corresponding interatomic distances (d_{ij}) for each BTS binding-poses from the obtained docking complexes [32].

4.2.4 Perturbation response maps and collectivity degree-based elastic network models. This computational approach evaluates the degree of change (*i.e.*, interatomic distances perturbation between residue fluctuations) induced by a given ligand (BTS) in the residues network (fibrinogen E-region) by describing the receptor-ligand complex ($fib_{(j)}$-$Lig_{(i)}$ *or R-L*) interaction potential (*U*) as a Hookean potential based on an elastic network analysis (ENM) [33]. Herein, the perturbation response-induced by BTS on fibrinogen binding-site was estimated by averaging across all the fibrinogen binding-residues for different amplitudes between the perturbed and the unperturbed state of the displacements from equilibrium using anisotropic vibrational analysis [33]. That is, using the following equation:

$$e_{p_{(i)}} = \frac{1}{N_a}\sum_{j=1}^{N_a}|p_j - u_j| \quad (1)$$

where $e_{p_{(i)}}$ stands for the effect of the perturbation response in the anisotropic normal mode (*i*), p_j is the displacement of the fibrinogen E-region residue (*j*) in the perturbed normal mode ($\left(d_{ij} - d_{ij}^0\right)^2$ R-L), u_j stands for the displacement of individual fibrinogen interacting residues (*j*) in the unperturbed normal mode ($\left(d_{ij} - d_{ij}^0\right)^2$ R-R), and N_a is the number of binding site residues obtained from the 3D Lig-Plot diagram analysis of each BTS binding-poses in the fibrinogen E-region protein receptor. The $e_{p_{(i)}}$ effect is depicted as a local perturbation response scanning map (LPRS map) for the different BTS binding-poses. Here the i^{th} rows are related to the response generated upon perturbing fibrinogen residue (*i*) and its average (*i.e.*, over all *k*-receivers residues of fibrinogen). While the k^{th} columns of the LPRS maps depict the sensitivity in response to the perturbation for all the fibrinogen allosteric residues (*j*-

sensors residues) [43]. We also consider other recognized parameters, like the degree of collectivity (K_k) for a given *k*-normal mode. This biophysical parameter informs about how the structural elements (fibrinogen binding residues) move together in that particular mode (*k*). Then, K_k is useful to estimate the cooperativity degree between residues in a given *k*-normal mode as shown below.

$$K_{k(R-L)} = \frac{1}{N} e^{-\sum_i^N (\Delta d_{ij})^2 ln(\Delta d_{ij})^2} \quad (2)$$

where *k* is the normal mode number, Δd_{ij} (= $d_{ij} - d_{ij}^0$) stands for the amplitude of the individual displacement induced by the different BTS-ligand binding poses (*i.e.*, displacements from R-L perturbation) as previously defined [42,43].

4.2.5 LPRS-maps image analysis-based on Fractal and Fast Fourier Transform. Herein, we introduced for the first time a combination of *Fast Fourier Transform* (2D-FFT) and two fractal approaches – that is, Mandelbrot set and box-counting methods [44-48], in order to identify different perturbation signals and fractal patterns based on local geometrical perturbations from the computational image processing. Those images were obtained from the LPRS maps for the different BTS-binding poses interacting with the fibrinogen E-region. For this purpose, we applied digital image processing based on 2D-Fast Fourier Transform (2D-FFT) for the previously obtained unperturbed and perturbed fibrinogen LPRS-maps. The FFT algorithm is a powerful digital image processing tool, which is used to decompose an image (LPRS maps) into its *sine* and *cosine* components. The number of frequencies ($\omega = 2\pi kn/N$) corresponds to the number of pixels (*N*) in the original LPRS map image (input image). These calculated frequencies contain relevant information about the geometric pattern of the biological response [44-48]. This fact allows the identification of spatiotemporal properties and patterns of complex pharmacological systems, such as the ones here targeted, *i.e.* the BTS-ligand binding conformations in the fibrinogen protein receptor. More details on this modeling algorithms will be discussed in the next section. Following the same objective, we apply the computational LPRS maps image modeling based on Mandelbrot set (or ***M***), and found to exhibit an infinitely complicated boundary that shows progressively ever-finer recursive details according to the increasing image magnifications [44-52]. The geometry shape of this repeating detail depends on the region of the set being studied in the LPRS maps.The Mandelbrot set is formed for an infinite number of distorted copies of itself and the central region of any of these copies represents an approximate cardioid. These cardioids structures

are generated by the rolling trajectory described by a point (which represents a given perturbed or unperturbed fibrinogen-residue) included inside a *circle 1* around a fixed *circle 2* with the same radius. In addition, the Mandelbrot set is linked to a Julia set, which generates self-similarly complex geometries based on fractal shapes [44-52]. To generate the Mandelbrot set images for the different LPRS maps of BTS-ligand binding modes, we used the GNU Image Manipulation Program_GIMP software. By the other hand, we applied the *Box counting method*, carried out by sampling image-data like fractal dimensions (FDs by computing the number of binary black (*B*) and white (*W*) pixels or boxes representing different types of FD, namely: D_{BW}, D_{BBW} and D_{WBW}[48-52].

4.2.6 Performing experimental methods

4.2.6.1 Materials. Bovine plasma fibrinogen, fraction I, type IV, was obtained from Sigma (No. 9001-32-5). 3-(2-Benzothiazolylthio)-propanesulfonic acid (BTS) (> 97% purity) was purchased from Aldrich Lancaster. The buffer solution used for fibrinogen was 50 mN glycine plus sodium hydroxide to give a pH of 8.5 (fibrinogen will be negatively charged). Samples were prepared within one hour prior to usage. All chemical reagents were of analytical grade, and solutions were made using doubly distilled and degassed water.

4.2.6.2 Isothermal titration calorimetry. Experiments were carried out in a VP-ITC microcalorimeter (MicroCal Inc., Northampton, U.S. [53]. In order to determine the binding isotherms, the drug solution 1 mM were introduced into the syringe (296 μL), while the fibrinogen solution 0.033 mM were introduced into the sample cell (1.4166 mL). The stirring was kept constant at 416 rpm. An important factor to take into account is the equilibrium time necessary before starting each experiment, in our case it was one hour, more than enough for the power base line to remain stable. 28 injections of 5 μL at a constant rate of 0.5 μL s^{-1} plus a first one of 2 μL were introduced in the cell every 300 s. To eliminate negative signals, a reference power of 25 μJ s^{-1} was applied. Thus, it is guaranteed that the signal is not altered by any overcompensation mechanism. Dilution experiments of pure drugs were also conducted. The values obtained were systematically subtracted from those measured for the fibrinogen-drug systems. In this way, it is guaranteed that all the heat produced in the cell is due solely to the binding process. Following the described procedure, experiments were performed at a temperature of 298.15 K.

4.2.6.3 SAXS acquisition and data analysis. SAXS experiments were carried out at SAXS beamline from the National Synchrotron Light Laboratory (LNLS) Campinas, Brazil. SAXS experiments from BTS at concentrations of 70, 110, 150 or 180 mM (cmc around 170 mM, at pH 5.5 at room temperature) were taken in the absence and presence of 0.5 mg/mL of fibrinogen-containing buffer solution (50 mM glycine at pH 8.5). SAXS curve from 0.5 mg/mL of fibrinogen was also obtained. Samples were placed perpendicular to the incoming X-Ray beam (wavelength $\lambda = 1.488$ Å) in between a mica window. The scattered X-Ray data were recorded by a 2D Pilatus detector in the q interval ranging from 0.012 $Å^{-1}$ to 0.45 $Å^{-1}$, where q is the scattering vector ($q = (4\pi/\lambda)\sin\theta$ with 2θ the scattering angle). All measurements were performed at room temperature of 298 K[54-59].

The scattering intensity $I(q)$ is related to the scattering particle form factor, $P(q)$, and structure factor, $S(q)$, as described by the equation 3:

$$I(q) = kn_pP(q)S(q) \quad (3)$$

where k is an instrumental parameter and n_p is the scattering particle number density [54-59]. A qualitative analysis of the SAXS curves from fibrinogen in the presence of a given BTS concentration was done by comparing the experimental data with the expected ones considered as a linear combination of the individual scatterings from the fibrinogen and BTS in solution.

4.3 Results and discussion

4.3.1 Theoretical section

4.3.2 Molecular docking and 3D Lig-Plot diagrams by BTS-binding poses. An important step to ensure accuracy of our theoretical approach is the identification/prediction of feasible binding-sites in the receptor protein molecule. In the present study the prediction of fibrinogen-binding sites was performed by using an *ezPocket* tool with maximum score equal to 1.0 for the fibrinogen molecule. This binding pocket detection algorithm works with a consensus approach using the *fconv* method [28,29] which is based on Delaunay triangulation to detect plausible fibrinogen binding pockets, typically taking a few seconds with accurate Cartesian *XYZ*-coordinates for the docking box simulations. In our case, this modeling task acquires a paramount importance because we tackle the docking study with a single molecule (BTS). The different binding-conformations that BTS can adopt into the fibrinogen-binding site can significantly influence its Gibbs free energy of binding (FEB) following two criteria: **1)** by inducing a little studied phenomenon so-called *intrinsic ligand-promiscuity* [60,61] according to the number of evaluated conformations in the same biophysical environment of a given protein receptor and **2)** attributable to the proteins having at least one similar binding-

cavity that could, in principle, accommodate small ligands in different conformations. Based on this, we can suggest that the simultaneous modulation of several target residues (effector and allosteric residues) by means of one single BTS molecule can be evaluated in a robust and unequivocal way and also with different types of non-covalent docking interactions, such as hydrophobic, H-bond, cation-Π, and Π-Π stacking interactions with its corresponding interatomic distances (d_{ij}) for each BTS binding-poses from the obtained docking complexes [62]. **Figures 1** and **SM1** (the latter shown in the Supporting Information – SI) illustrate and compare the different BTS binding-poses.

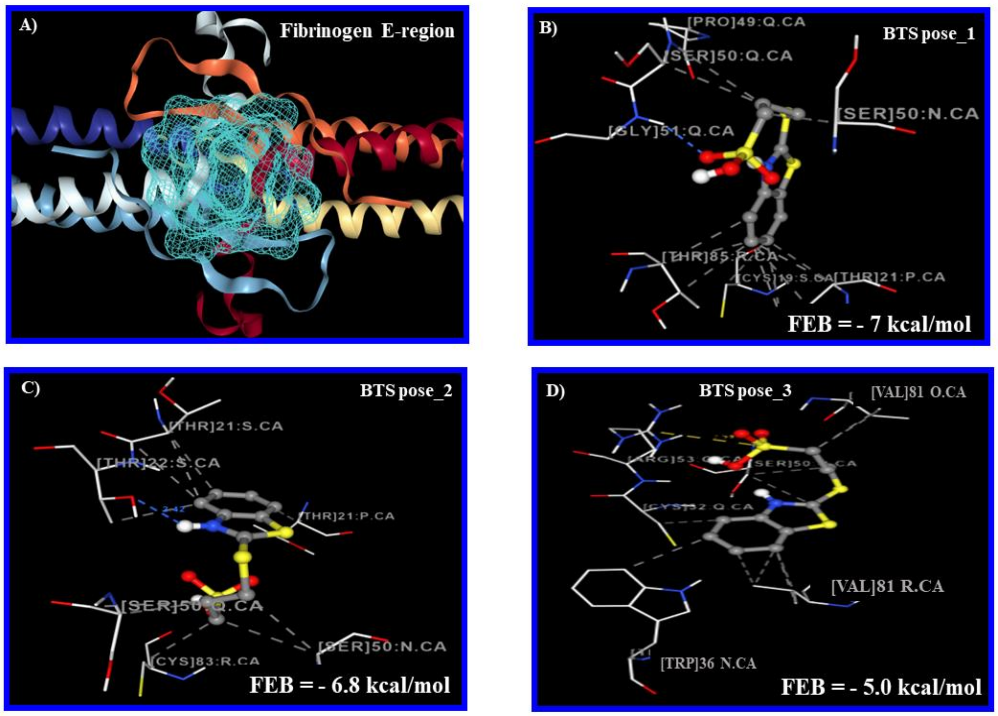


Figure 1. Results of molecular docking simulation between BTS_binding poses and fibrinogen E-region. **A)** Theoretical prediction of the main fibrinogen E-region binding-pocket generated by *ezPocket* tool with the *fconv* method [28,29]. The 3D-lig-plot diagrams for best BTS binding-poses are shown like **B)** BTS_*pose 1*, **C)** BTS_*pose 2* and **D)** BTS_*pose 3* obtained from the docking complexes BTS-fibrinogen, and depicted with their corresponding FEB values in kcal/mol. Herein, gray dotted-lines between the different BTS-poses and the fibrinogen residue interacting atoms represent van der Waals hydrophobic interactions and the blue dotted-lines represent H-bond contacts. The remaining cases like BTS_poses (from 4 to 9) can be found in **Figure S1**. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

It is important to note that, the formation of the docking complexes follows a spontaneous thermodynamic process (FEB < 0 kcal/mol) for the nine conformations evaluated with a high prevalence of hydrophobic interactions in most cases and with different patterns of interacting fibrinogen residues. By the other hand, the presence of H-bond contacts contributes to the stabilization of the formed BTS-fibrinogen complexes, but these were only identified for the BTS_*pose 1* (FEB = −7.0 kcal/mol), BTS_*pose 2* (FEB = −6.8 kcal/mol), BTS_*pose 4* (FEB = −4.8 kcal/mol), and the BTS_*pose 5* (FEB = −4.7 kcal/mol), which in turn present the most negative FEB values of interactions according to more stable docking complexes.

4.3.3 Perturbation response maps and collectivity degree by BTS-ligand conformation. Let us now discusss the relationships between the different BTS-docking poses, their ability to induce local-perturbations in the fibrinogen residue network, and the change in the collectivity degree [63-67]. The local perturbation response scanning maps (LPRS-maps) show the different patterns of perturbation (interactions) with the fibrinogen E-region target-residues in the absence (unoccupied fibrinogen E-region) and presence of different conformations adopted by BTS. Given a protein, it is well-known that vibration modes between its residue pairs can be affected by the presence of ligands due to changes induced by residues' distance fluctuations and deformation energy effects [63-67]. In fact, the transitions between the functional states in the fibrinogen E-region binding-site are strongly dependent on the fibrinogen flexibility properties owing to the presence of allosteric inter-communications between cluster of residues. In this respect, we theoretically suggest that, the different BTS_*poses* evaluated can directly affect the residue network topology (*i.e.* the interaction matrix) for effector (*i*) and allosteric residues (*j*) of the fibrinogen binding-site. The results obtained from this perturbation response analysis are illustrated in **Figures 2** and **SM2** (SI). The i^{th} rows LPRS-maps are linked to the effector response triggered upon perturbing residue (*i*) and to its average (*i.e.*, over all receiver residues of fibrinogen). Likewise, the k^{th}-columns of the LPRS-maps show the sensitivity in response to the perturbation induced by BTS-ligand poses in all the fibrinogen allosteric residues (*j*-sensors residues).

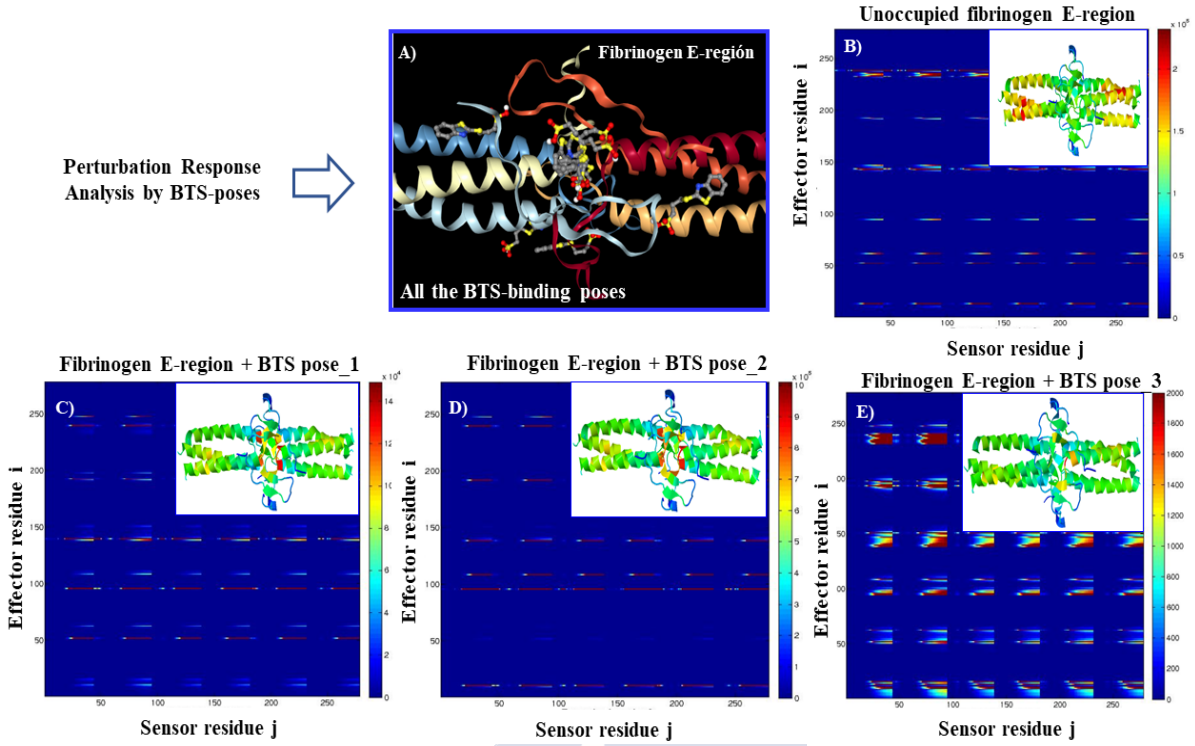


Figure 2. Perturbation response analysis for the BTS-binding poses. **A)** 3D-Cartoon representation for all BTS-docking poses superimposed in the fibrinogen E-region binding site, designed using the open-source Pymol visualization system. **B)** LPRS map obtained for the unoccupied fibrinogen E-region binding site and acting as the control simulation experiment. Individual LPRS-maps obtained from the best BTS-fibrinogen docking complexes-based conformation analysis and the corresponding 3D-fibrinogen E-region structure (for all modes) with intensity bar color (on the right) for: **C)** BTS_*pose 1*, **D)** BTS_*pose 2*, and **E)** BTS_*pose 3*. For this purpose, dark blue to red regions represent the increasing propensity to act as sensor/effector after the stimulus of perturbation (BTS-interaction) for 278 residues analyzed from the fibrinogen E-region like by C(α)-atoms. All the LPRS-maps were established in range of the low frequency normal modes equal to 2. The remaining cases like BTS_poses (from 4 to 9) can be found in the **Figure S2**. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

Based on these LPRS maps, we found several differences on the behavior of the biophysical parameter collectivity degree $K(k)$ [63-67]. That is, in general terms, we discovered a slight tendency to decrease when we compare the $K(k)$ values for the case of BTS-binding poses (BTS_*pose 3-5*), which are around ~ 0.5 in the normal mode 2 (**Figures 3** and **SM3**), with the ones for the unoccupied fibrinogen E-region. A high $K(k)$

value is frequently associated to high entropy, which means that highly cooperative normal modes are influenced by a significant region or domain in the protein structure, *i.e.*, the ligands show to induce perturbations with motion fluctuations distributed over a larger number of residues that can be extended to the full protein. From the biophysical point of view, a higher $K(k)$ can trigger cooperative allosteric responses that are linked to lower frequency of motion of the protein. This is mainly of interest to explain the protein dynamics and physiological functions when the changes in docking free energy (FEB values) are imperceptible [63-67] like those induced by the different binding-conformation (BTS_poses *1-9*) with respect to the reference simulation condition – unoccupied fibrinogen E-region used as control experiment.

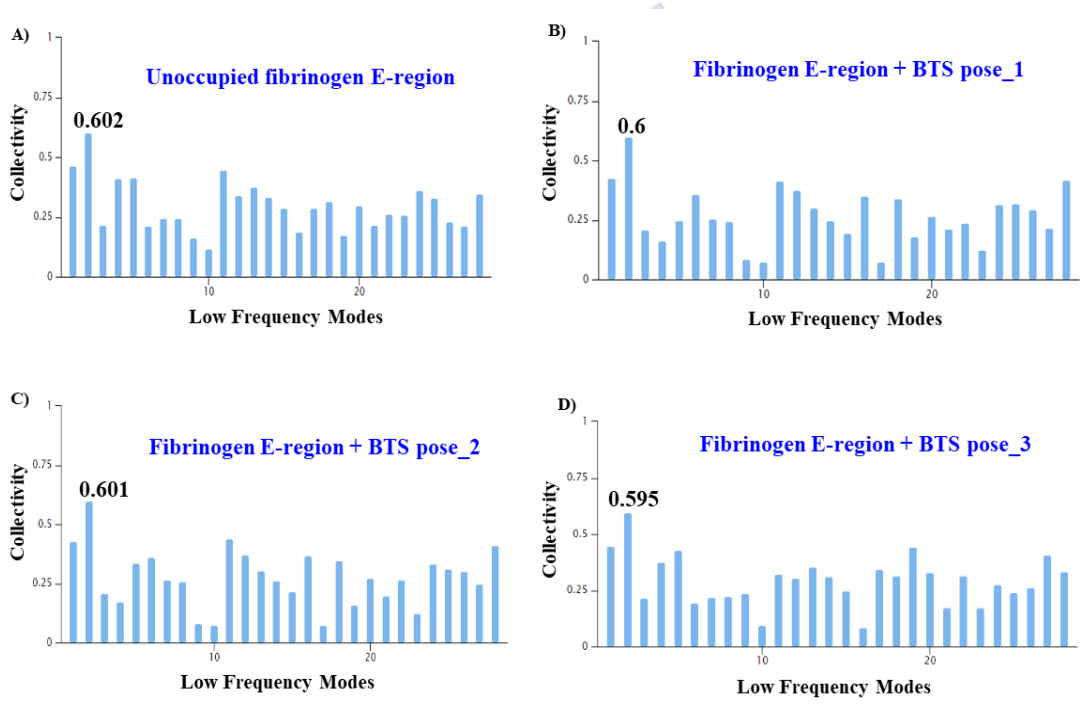


Figure 3. Graphical representation of the degree of collectivity $K(k)$ *vs.* low frequency normal modes (mode 2) for the control **A)** unoccupied fibrinogen E-region (0.602), and for the three best BTS-binding poses, namely: **B)** Fibrinogen + BTS_*pose 1* (0.6), **C)** Fibrinogen + BTS_*pose 2* (0.61), and **D)** Fibrinogen + BTS_*pose 3* (0.59). The remaining cases can be visualized in **Figure S3**. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

4.3.4 Fourier transform based on BTS-ligand conformational analysis. Herein, we applied 2D-FFT approach to explain non-trivial interaction patterns based on the different conformations of BTS ligand with the fibrinogen protein receptor [9, 68-72]. To do so,

we gather the individual profiles coming from the 2D-FFT analysis of the nine conformations adopted by the BTS ligand in the fibrinogen E-region binding-site, and resulting from the computational image processing of the local perturbation response maps of BTS-ligand conformations. The results of this approach are presented in **Figure 4**.

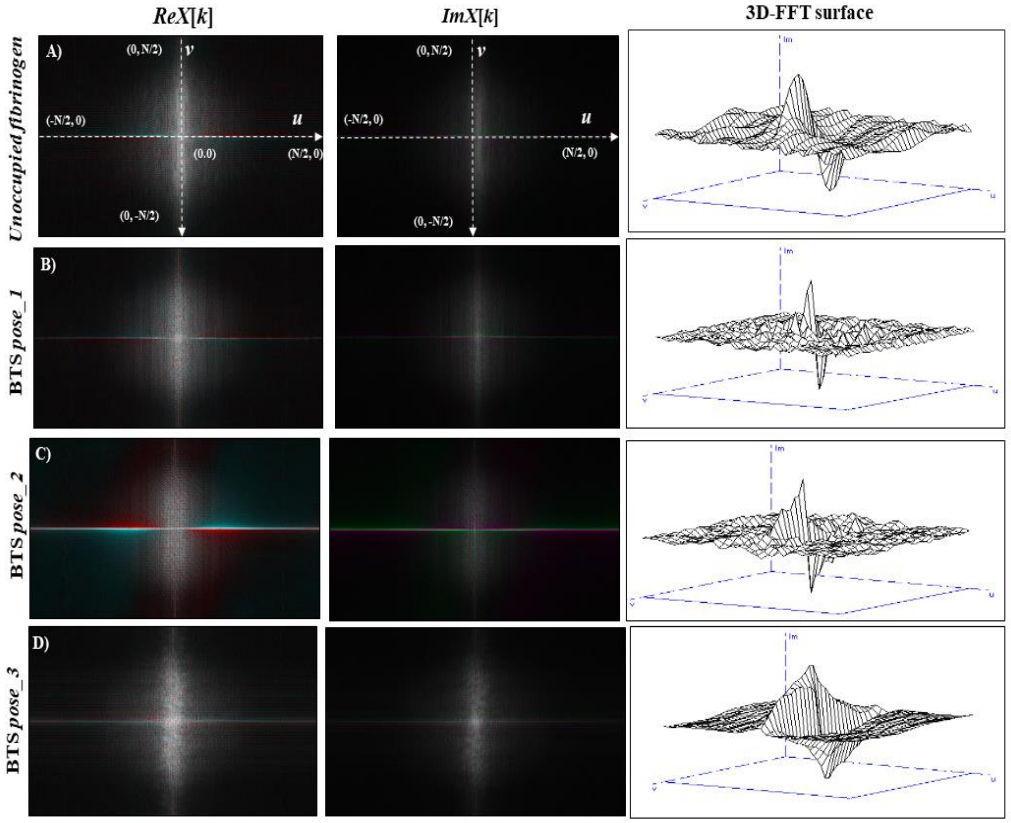


Figure 4. Graphical representation of the 2D-FFT profiles of the frequency domain ($X[k]$) based on the real component ($ReX[k]$) and its imaginary component ($ImX[k]$), as well as the 3D-FFT surface graph with $ImX[k] \neq 0$ and $ReX[k] = 0$ obtained from the LPRS-maps for the control **A)** unoccupied fibrinogen, and for the three best BTS-crystallographic conformational binding modes, *i.e.*: **B)** fibrinogen plus the BTS_*pose 1*, **C)** fibrinogen plus the BTS_*pose 2*, and **D)** fibrinogen plus the BTS_*pose 3*. For all the 3D-FFT surface graphs, the coordinates axes (u, v) are included into the spatial frequency to obtain the corresponding signal 3D-frequency domain ($X[k]$) based on the imaginary dimension, the latter representing the signal perturbation-response (BTS pharmacodynamic interactions) depicted in the complex plane k from a qualitative point of view. The remaining cases can be visualized in **Figure S4**. Source Elsevier article,

please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

Specifically, the LPRS-map images were decomposed to a linear combination of two harmonic orthogonal functions to obtain the symmetric 2D-FFT spectrum, in which the lowest frequency (*i.e.*, the average pixel values) is depicted near the center [0,0], whereas the highest one can be detected near the corner of the 2D-FFT spectrum. The differences between the 2D-FFT patterns obtained from each BTS-ligand conformation are established by the direction of the lines in the spectra, which in turn are defined by the direction of the 2D-spatial sinusoids. For the BTS-binding poses up to 9, we show the 2D-FFT frequency spectra coming from the filtration of 2D-perturbation signals of the respective LPRS-maps, with zero frequency at the center. That is, the line from each pair of points (pixels) was drawn at the center of the 2D-FFT image. This signal spectra representation is simpler to understand, since all its lines are drawn from the same point. Besides the advantage of placing the zero at the center it is the fact that, it fits with continuous frequency spectra signals of the images that are consistent with the modeling principles for complex biological systems, and allow the identification of frequency (spectral) content analysis of a signal just like ligand-protein interaction perturbation-signals. When the input image is continuous, then its signals are aperiodic. This fact places zero frequency at the center of the 2D-FFT frequency spectrum, with the $X[k]$ signals becoming higher in all directions out to infinity [73-75]. As far as our patterns are concerned, the zero frequency components (central peaks) are the largest. This is not surprising because most of the images from nature are in the low frequency range. Also patterns containing vertical (more intense) and horizontal lines (less intense) indicate the existence of components at all frequencies, but their intensity decreases with the frequency. This means that most information is placed in the low frequencies, that is, the more important contributions arise from slow changes in LRSP maps. So, conformational changes spread smoothly along the polypeptide chains. These changes are mostly one-dimensional as the vertical dominant direction suggests. The absence of rotations in the line indicates that none of the BTS-poses involve a rotational movement of the E-domain. The most intense 2D-FFT central signals were detected for BTS_*pose 1*, BTS_*pose 2* and BTS_*pose 3* compared to the reference control. For the other BTS-poses, the central peak decreases and higher frequency contributions increase.

4.3.5 Fractal analysis-based on LPRS-maps of BTS-ligand conformation. Lastly, we examine the theoretical results following a combination of two recognized algorithms for fractal analysis: **(i)** the Mandelbrot set and **(ii)** box-counting methods [45, 47, 51]. The key idea behind such approach is the concept of the fractal protein, through which, the ligand-druggability properties of a simple protein (fibrinogen E-region binding-site) is formally represented as a highly complex fractal form according to non-Euclidean geometry, irregularity properties of the protein surface and the principles of fractal theory [45, 47, 51].

In this context, the main goal of this approach was to describe the differences in the fractal geometric-pattern of the perturbation responses in the fibrinogen binding-site, derived from the LPRS maps before and after the BTS-docking interactions, from a qualitative and quantitative point of view [70]. From the mathematical point of view, the fractal protein (fibrogen) is a finite square subset of the Mandelbrot fractal set ($\boldsymbol{M}$), with center coordinates (x, y) and sides of length z, defined by a family of complex quadratic polynomials function (f) obtained by using an Iterated Function System (IFS) that is executed in a feedback loop or quadratic recurrence equation given by $\boldsymbol{f_c(z)} = \mathbf{z_{n+1} = z^2 + c}$ [46-50]. Due to this, it is possible to achieve a high degree of complexity based on the inter-residue network communication derived from its folding properties (roughness), which is more significant in the surfaces of interaction for most of the catalytic binding-sites of the protein (fibrinogen) being able to be disturbed in the presence of certain ligands (BTS).

For this instance, the $\boldsymbol{M}$ set can be defined as the set of all the N-pixels of a given LPRS map image, which represent an unperturbed or perturbed fibrinogen binding-residue in such a way that the previous sequence $\boldsymbol{f_c(z)}$ does not escape to infinity. Then, a complex number $\boldsymbol{c}$ is part of the Mandelbrot set $\boldsymbol{M}$ if, when applying the quadratic recurrence equation, the iterating function $fc\mathbf{(z)}$ remains bounded. For the first time, the initial value starting with $\mathbf{z_n = 0}$ is taken as a critical point like $f_c^n(0)$. Then, we can formally define the Mandelbrot set $\boldsymbol{M}$ as follows:

$$\boldsymbol{M_{set}} = \boldsymbol{f_c^n(z_n)} = \lim_{n\to\infty} \|z_{n+1} = z_n^2 + c\| \nrightarrow \infty \text{ where } z_n = 0 \qquad (4)$$

Here, the IFS output (recursive function) of $\boldsymbol{f_c^n(z_n)}$ is taken as input when the function calls itself and the absolute value of $\boldsymbol{z_n}$ it remains bounded. The vertical brackets denote

the Euclidean norm, which is a measurement of how far away a point in the complex plane (k) is from origin: $\|z\| = \sqrt{a^2 + b^2}$. The results obtained by applying the Mandelbrot set approach are presented in **Figure 5**.

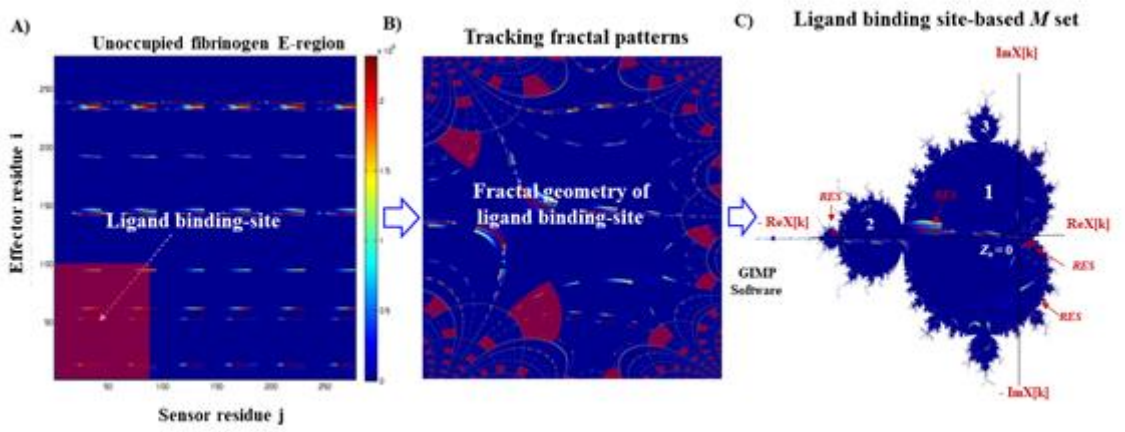


Figure 5. **A)** LPRS map for unoccupied fibrinogen (control) highlighting the fibrinogen binding-site (*i.e.*, the interacting residues from E-region are included in the transparent red rectangle). **B)** Data visualization-based tracking fractal patterns of the fibrinogen binding-site. **C)** Data visualization-based Mandelbrot set fractal approach with cardioid image modeling obtained for unoccupied fibrinogen binding-site with different positions for the ligand binding-site residue cluster (*RES*) in the relevant Mandelbrot cardioid (*i.e.*, cardioids 1, 2) considered like unperturbed fibrinogen residues (*RES-unbound*). The Mandelbrot set parameters are fixed at $ReX[k]_{max} = +1$, $ReX[k]_{min} = -2$, $ImX[k]_{max} = +1$, and $ImX[k]_{min} = -1$ for comparison purposes with the different BTS-poses. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

These Mandelbrot set results refer to the protein-ligand binding condition for the three best BTS-docking *poses_1*-3, in order to identify critical changes in the fractal geometry of the fibrinogen binding residues (*i*-effector and *j*-sensor residues; *RES*) before and after the interaction with the different BTS-binding_*poses* (see **Figure 6**).

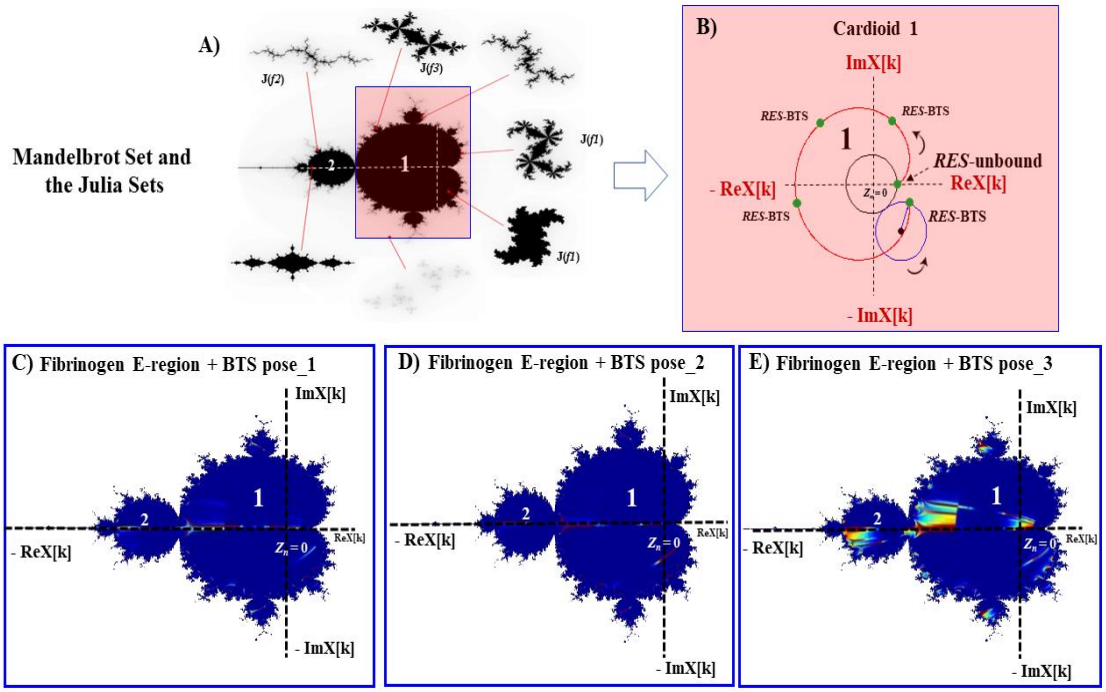


Figure 6. Data visualization subsets of the Mandelbrot set fractal approach with cardioid image modeling from LPRS maps. **A)** Canonical representation of the ***M*** set with the main cardioids (1 and 2) and the corresponding Julia sets ($J(f2)$, $J(f3)$ and $J(f4)$). **B)** Geometrical representation of the different positions of a given fibrinogen residue cluster (*RES;* green dot) in the main Mandelbrot cardioid (*i.e.*, cardioid 1) considered like an unperturbed fibrinogen binding-site residue cluster (*RES-unbound*) in $z_n = 0$ before BTS-interaction, and the same perturbed fibrinogen residue cluster (*RES-BTS*) after BTS-interaction in the complex plane *k*. In addition, the data visualization subsets of the Mandelbrot set fractal approach for the three best BTS-ligand binding modes are shown. **C)** Fibrinogen + BTS_*pose 1*, **D)** Fibrinogen + BTS_*pose 2*, **E)** Fibrinogen + BTS_*pose 3*. The Mandelbrot set parameters are fixed at $ReX[k]_{max} = +1$, $ReX[k]_{min} = -2$, $ImX[k]_{max} = +1$, and $ImX[k]_{min} = -1$. The remaining cases can be found in **Figure S5**. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

Let us begin by looking at the fractal patterns in the *n*-cycle components attached within the big region of the Mandelbrot set (*i.e.*, the main *cardioid 1*) and around it. The cardioid image modeling-based Mandelbrot set is defined as the trajectory described by a point (*i.e.* the LPRS map image pixel) of a circle that rolls around another fixed circle with the same radius to get a modified or unmodified cardioid image pixel, which represent

differences in the fractal geometry shape of the response (*i.e.* the docking interactions coming from the LPRS maps). That is to say, it represents the trajectory described by any relevant interacting fibrinogen-residues (*RES_fib*) before and after their interactions with the BTS-poses up to 9. For example for the *RES_fib-BTS_pose 1*, these are: PRO49_chain Q, SER50_chain Q, TRH85_chain R, CYS19_chain S, and for the *RES_fib-BTS_pose 2*: TRH21_chain S, TRH22_chain S, SER50_chain Q, CYS83_chain R, SER50_chain N), which together make up the protein binding pocket (Please, refer to **Figure 1. B** and **C**.). Actually, the main *cardioid 1* comprises a relevant region of geometric parameters for which it has an attracting fixed point with period $n = 1$. The obtained subsets of the Mandelbrot set (cardioid 1) thus provide relevant information about potential structural changes associated with the fractal geometry of the protein-ligand systems evaluated (unoccupied fibrinogen E-region and BTS-poses at the fibrinogen E-region). Indeed, previous works have suggested that changes in the spatial ordering of protein binding residues induced by certain stimuli (like ligands) can modulate the binding properties of the protein through changes in its fractal dimension (FD) [75,76]. It is well known that the FD is associated to the backbone non-Euclidean geometry, as well as to the irregular nature and fractal surface properties of the binding sites. This is justified by the fact that most of the ligand-protein interaction processes occur under strict conditions of ligand-specificity and, at the same time, that these processes depend on surface phenomena with a defined geometric pattern of complementarity. We thus suggest that small changes in the fractal surface patterns, such as the cardioid 1 subset, could affect not only the native fibrinogen folding and solvent accessible surface in the unbound state (unoccupied fibrinogen E-region), but also the conformational entropy and thermodynamic stability of the docking complexes formed between the fibrinogen E-region and the different ligands [75,76]. The major contribution to these changes comes from the establishement of van der Waals and H-bond interactions mainly with fibrinogen *i*-effector residues from the chains N, Q, R, and S. It is therefore easy to identify dissimilarities between the fractal patterns-based subsets of the Mandelbrot set for the best BTS conformations evaluated (BTS-binding poses_1-3), which are mainly concentrated in the ***M*** region (*cardioid 1*) in its positive (*ImX*[k]) and negative (*ReX*[k]) quadrant, predicting changes in the geometric complementarity of the *i*-effector residues from chains N, Q, R, and S. In addition, the changes can be confronted with those from the Mandelbrot set of the control (unoccupied fibrinogen E-region) in the same quadrant. By the other hand, the differences observed between the best BTS-binding poses_(*1-3)* evaluated and the unoccupied fibrinogen E-

region in the small quasi circular-shaped region, named *cardioid 2*, which is located in the left-site of *cardioid 1* with negative real and imaginary components, $-ReX[k]$ and $-ImX[k]$, respectively. The fractal-geometric differences in *cardioid 2* are maybe associated with changes in the geometric parameters for the clusters of *j*-sensor residues (allosteric residues), for which the attracting cycle has a period of 2 following complex geometrical criteria. Furthermore, other different fractal-patterns were detected for an infinitely many other periodic components like small *cardioids* tangent to *cardioid 1* and *cardioid 2*, such as ***M***-submaps of *n*-cycles (called *hyperbolic components*) that exhibit a "*chaotic*" behavior, plotted in the interior regions of ***M*** and over *Julia sets J*(*f*) [47-49].

In order to quantify potential fractal geometrical perturbations between the formed docking BTS complexes and the fibrinogen binding-site, we calculated their fractal dimensions by using the box-counting algorithm [51]. Notice that one of the most important properties in the fractal analysis is the degree of self-similarity. For instance, a topological fractal dimension near to 2 is considered to embody high complexity (high variety of geometrical information) and low self-similarity, while a topological fractal dimension closer to 1 represents little complexity and high self-similarity [47-49]. Following these criteria, the LPRS-maps were analyzed using the box counting algorithm and the obtained results are shown in **Figure 7**. For this instance, the non-Euclidean geometrical patterns were characterized according to the fractal dimension like D_{BW}, which describes the LPRS maps properties by considering the border of LPRS map fractal pattern. $D_{BBW,}$ characterizes the LPRS map-fractal pattern on the white background, and the D_{WBW} the LPRS map-fractal pattern on the black background of the images calculated for each BTS-*poses_1-9*.

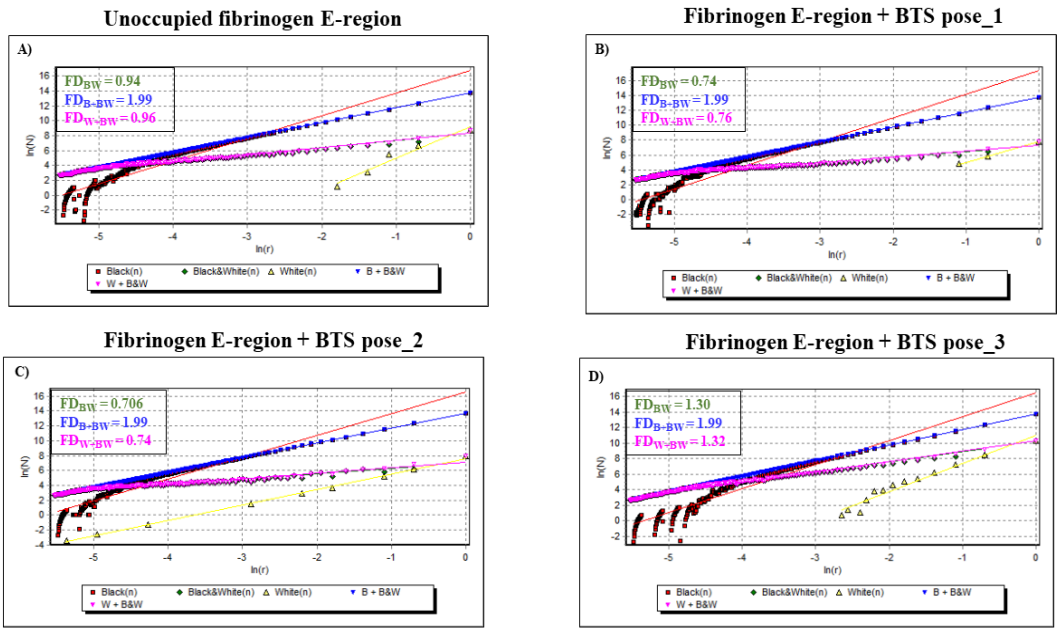


Figure 7. Fractal spectrum based on the box-counting algorithm applied to obtain the slopes of the linear regression yields from binary black/white LPRS maps image-processing. These slopes correspond to the fractal dimensions (*FD*: D_{BW}, D_{B+BW}, and D_{W+BW}) for the **A)** unoccupied fibrinogen binding-site and the best docking complexes like: **B)** fibrinogen + BTS_*pose 1*, **C)** fibrinogen + BTS_*pose 2*, and **D)** fibrinogen + BTS_*pose 2*. The remaining cases can be seen in **Figure S6**. Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

The calculated FDs are associated to the fibrinogen surface and backbone non-Euclidean geometry. These can show how the protein folding, packing density, solvent-accessibility, crystallographic positioning and binding-interaction properties can be affected under the influence of different conformation of the BTS-ligand in the fibrinogen binding-site. Especially, by the number of van der Waals docking-interactions through the different BTS-binding modes, degree of collectivity $K(k)$, and van der Waals distances of interaction (≤ 7 Å). Herein we suggest that, BTS-ligand binding modes lead to a change in roughness-based FDs [75,76]. Because we detect that the fractal dimensions (D_{BW} ~1.05 and D_{W+BW} ~1.07) were higher for all the BTS-*poses_1-9* complexes at the fibrinogen E-region compared to the physiological condition of unoccupied fibrinogen E-region binding site used as control (D_{BW} = 0.94 and D_{W+BW} = 0.96, respectively). As mentioned above, a FD (D_{W+BW}) ≈ 2 reveals high variety of geometrical information and

low self-similarity, while FD (D_{W+BW}) ≈ 1 represents little complexity and high self-similarity. It is important to note here that, the fibrinogen protein is a large macromolecule and its vibrations generate a huge amount of different conformational sub-states (normal modes-based degree of collectivity) under BTS-ligand interaction, which can be associated with local perturbations (*i.e.*, variation in the local magnitudes of calculated FD from the fibrinogen + BTS complexes).

By the other hand, the value of the D_{B+BW} remained fixed around 1.99 in all the cases, thus revealing high complexity of geometrical information. The latter suggest that the BTS-ligand can induce significant changes in the geometrical selectivity-based FD of the fibrinogen E-region binding site. It is well-known that changes (*global* and *local perturbations-induced by ligands*) in the spatial arrangement of atoms in key residues (like perturbations-induced on *j*-allosteric residues) of proteins can be described by their fractal dimension. Besides these changes can negatively impact the biological function of proteins, particularly when they affect the structural properties of the binding cavities, which are essential in the complementary processes of ligand geometrical specificity and recognition [72-75]. However, we note that the values of the D_{BW} and D_{W+BW} are very close to each other for the different BTS-poses complexes at the fibrinogen E-region. This in turn might be related with the very close strength of the Gibbs free binding energies (FEB values) obtained for most of the evaluated cases. The highest values for FD were found for the system BTS *pose_3* at the fibrinogen E-region (D_{BW} = 1.30 and D_{W+BW} = 1.32), which presents an average FEB = –5 kcal/mol. In this case, we suggest that the fractal structure of the local perturbation response maps obtained for the BTS *poses_1-9* exhibit low variety of geometrical information and high statistical self-similarity among them. In this regard, a fractal is considered as rigorously self-similar, if it can be expressed as a union of sets, each of which is an exactly reduced copy of the full set of points, taken as perturbed and unperturbed fibrinogen binding site residues (E-region). Additional details on black/white binary LPRS map imagen-processing can be found in **Figure S7**.

Results and discussion. Experimental validation

To further complete our computational results, we have performed two relevant validation experiments by using isothermal titration calorimetry (ITC) and small angle X-ray scattering (SAXS) experiments. The experimental results are presented in the **Figure 8.**

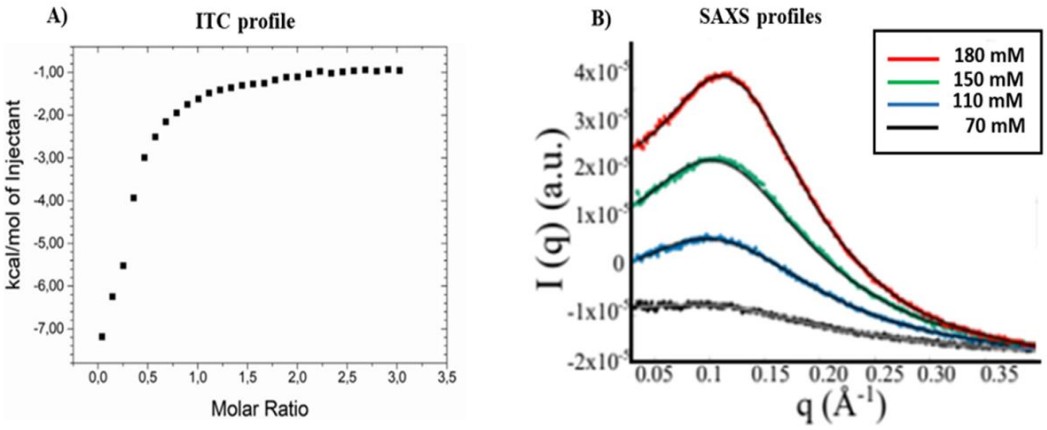


Figure 8. **A)** on the left; enthalpy as a function of molar ratio (BTS/fibrinogen). Herein, experiments were performed by adding 1 mM BTS solutions into 0.033 mM of fibrinogen. **B)** on the right, representation of experimental SAXS curves for fibrinogen protein in the presence of different BTS concentrations as 70 mM (black), 110 mM (blue), 150 mM (green), and 180 mM (red). The p(r) functions were calculated by using GNOM software. [77-79] Source Elsevier article, please visit: https://doi.org/10.1016/j.ijbiomac.2020.07.044 (accessed on 15 november 2020)

Firstly, we have explored the energetic contribution of BTS plus fibrinogen interactions by using ITC. For this instance, the corresponding microcalorimetric titration profile for the binding complex of BTS with fibrinogen was obtained (**Figure 8, A)**. Each point corresponds to the heat involved in a single injection of BTS (0.033 mM) into the solution containing fibrinogen (1 mM), corrected by subtracting the heats of mixing and dilutions. As depicted in the ITC profile an exothermic process with saturation occurring at BTS/fibrinogen molar ratio higher than 1 was obtained. Regarding the kinetic evolution of the fibrinogen/BTS complex, the equilibrium was reached after 150 minutes (obtained from calorimetric raw data). Then, experimental points were fitted following to the independent binding model in order to achieved the relevant thermodynamic parameters of the interactions process. Besides, other BTS/fibrinogen binding models were also checked but large uncertainties were obtained for the associated thermodynamics parameters. For this instance, the obtained values were: 0.98, 6.27×10^{4} M^{-1}, -7.30 kcal/mol, and -2.55 $cal/mol^{-1}{}^{o}C^{-1}$ for stoichiometry, binding constant, enthalpy and entropy, respectively. The stoichiometry reveals that just one type of binding interaction takes place. Furthermore, the binding constant suggests that BTS is able to acts as a

moderate ligand. Being the obtained enthalpy and entropy negative values. Concerning to the enthalpy, this means that the binding process is mainly driven by the formation of non-covalent interactions (i.e., H-bond, electrostatics, π-π stacking interactions); in the case of negative entropy change, it could reveals hydration and loss of conformational freedom associated with the BTS/fibrinogen complex formation fitting well with previous experimental evidences on this topic[78]. In connection with the free binding energy, the value obtained from ITC is -6.54 kcal/mol, which fit very well with the predicted docking results, it is important to note that for the BTS pose_1 and BTS pose_2, both in the quantities (-7 and -6.8 kcal/mol, respectively) and in the presence of hydrogen bonds and van de Waals hydrophobic interactions. Amaral et al [69] highlighted the differences in the thermodynamic quantities when the ligand interacts with a loop (loop-in compound) or an helix (helix-binding compound). Being basically, that loop-in binding compounds are mainly enthalpically governed, with favorable or small entropic contributions. While, for the helix-binding compounds, a large favorable entropic contribution to binding, and most of them have a binding enthalpy penalty are predominant. Clearly, our results evidence that BTS is a loop-in compound. This is another interesting point because docking results for BTS pose_1 and pose_2 reveal interactions with residues places between the 5-turn (or 5-turnI) and NCap as represented in the Lig-Plot results (refer to **Figure 1**). On the other hand, entropy change upon BTS binding can arise from a reduction in translational and rotational degrees of freedom, an alteration of the conformational flexibility of the binding partners, and from the reorganization of their solvation shells upon binding. Having ruled out desolvation effects, we then focus on conformational changes. In our perturbation analysis we have calculated the vibrational modes and protein flexibility pre and following BTS binding. The obtained degree of collectivity of fibrinogen was 0.602 (mode 2), while for the BTS pose_1 and BTS pose_2 decrease slightly to 0.6 and 0.601, respectively. This subtle decrease also manifests itself in all the other analyzed modes. A high degree of collectivity means a highly cooperative mode that enrolls a portion of the protein. Thus, upon binding cooperativity decrease, and also conformational flexibility, become the main contribution to entropy change.

Next, in order to evalute the impact of the BTS ligand in the fibrinogen conformational binding mechanims, the SAXS experiments were carried out by using a BTS-containing solution at 70mM, 110mM, 150mM, and 180 mM[79]. (See **Figure 8, B**).

Herein, It is also important to remark that a transition in the fibrinogen aggregation process must take place between 70 mM BTS and 110 mM BTS, due to the quite different

scattering profiles at low q range exhibited from these two experiments. However, with the increase of drug concentration the SAXS curves are similar for $q < 0.03$ $Å^{-1}$. This gives us an insight that similar protein–BTS complexes coexist with BTS aggregates in 110 mM to 180 mM BTS-containing solution.

The obtained SAXS measurements confirm and validate not only the presence of binding conformational changes-induced by different BTS docking poses in fibrinogen structure; but also show a strong dependence on BTS concentration.

References

[1] T. Ando, J. Skolnick, Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion, Proc. Natl. Acad. Sci. 107 (2010) 18457–18462.

[2] J.M. Ruso, N. Hassan, Role of biomacromolecules in biomedical engineering, Curr. Top. Med. Chem. 18 (2018) 1171–1187.

[3] A. Lopes, S. Sacquin-Mora, V. Dimitrova, E. Laine, Y. Ponty, A. Carbone, Protein-protein interactions in a crowded environment: an analysis via cross-docking simulations and evolutionary information, PLoS Comput. Biol. 9 (2013), e1003369. .

[4] S. Akira, S. Uematsu, O. Takeuchi, Pathogen recognition and innate immunity, Cell 124 (2006) 783–801.

[5] E.A. Meyer, R.K. Castellano, F. Diederich, Interactions with aromatic rings in chemical and biological recognition, Angew. Chem. Int. Ed. 42 (2003) 1210–1250.

[6] M.W. Gonzalez, M.G. Kann, Chapter 4: protein interactions and disease, PLoS Comput. Biol. 8 (2012), e1002819-e1002819.

[7] J.R. Kenny, M.M. Liu, A.T. Chow, J.C. Earp, R. Evers, J.G. Slatter, D.D. Wang, L. Zhang, H. Zhou, Therapeutic protein drug-drug interactions: navigating the knowledge gaps highlights from the 2012 AAPS NBC Roundtable and IQ Consortium/FDA workshop, AAPS J. 15 (2013) 933–940.

[8] L. Di, C. Breen, R. Chambers, S.T. Eckley, R. Fricke, A. Ghosh, P. Harradine, J.C. Kalvass, S. Ho, C.A. Lee, P. Marathe, E.J. Perkins, M. Qian, S. Tse, Z. Yan, M.J. Zamek Gliszczynski, Industry perspective on contemporary protein-binding methodologies: considerations for regulatory drug-drug interaction and related guidelines on highly bound drugs, J. Pharm. Sci. 106 (2017) 3442–3452.

[9] A. Bratek-Skicki, P. Żeliszewska, J.M. Ruso, Fibrinogen: a journey into biotechnology, Soft Matter 12 (2016) 8639–8653.

[10] M.P. De Maat, Effects of diet, drugs, and genes on plasma fibrinogen levels, Ann. N. Y. Acad. Sci. 936 (2001) 509–521.
[11] A. Lodi, A. Saha, X. Lu, B. Wang, E. Sentandreu, M. Collins, M.G. Kolonin, J. DiGiovanni, S. Tiziani, Combinatorial treatment with natural compounds in prostate cancer inhibits prostate tumor growth and leads to key modulations of cancer cell metabolism, npj Precision Oncology 1 (2017) 18.
[12] M.H.M. Leung, T.W. Kee, Effective stabilization of curcumin by association to plasma proteins: human serum albumin and fibrinogen, Langmuir 25 (2009) 5773–5777.
[13] V.V. Kleandrova, F. Luan, H. González-Díaz, J.M. Ruso, A. Melo, A. Speck-Planche, M.N.D. Cordeiro, Computational ecotoxicology: simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions, Environ. Int. 73 (2014) 288–294.
[14] Z.J. Deng, M. Liang, I. Toth, M.J. Monteiro, R.F. Minchin, Molecular interaction of poly (acrylic acid) gold nanoparticles with human fibrinogen, ACS Nano 6 (2012) 8962–8969.
[15] Y. Liu, X. Tang, J. Pei, L. Zhang, F. Liu, K.a. Li, Gastrodin interaction with human fibrinogen: anticoagulant effects and binding studies, Chem. Eur. J. 12 (2006) 7807-7815.
[16] N. Hassan, P. Verdes, J.M. Ruso, Assessment of interactions between four proteins and benzothiazole derivatives by DSC and CD, J. Chem. Thermodyn. 43 (2011) 399–404.
[17] M. Subramanyam, R. Sreenivasulu, R. Gundla, M.V. Rao, K.P. Rao, Synthesis, biological evaluation and docking studies of 1, 3, 4-oxadiazole fused benzothiazole derivatives for anticancer drugs, Letters in Drug Design & Discovery 15 (2018) 1299–1307.
[18] R.S. Keri, M.R. Patil, S.A. Patil, S. Budagumpi, A comprehensive review in current developments of benzothiazole-based molecules in medicinal chemistry, Eur. J. Med. Chem. 89 (2015) 207–251.
[19] D.C. Henry, A. Lapworth, The cataphoresis of suspended particles. Part I. Tthe equation of cataphoresis, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 133 (1931) 106–129.
[20] J.G. Greener, M.J.E. Sternberg, AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis, BMC bioinformatics 16 (2015) 335.
[21] D.E. Mager, D.R. Abernethy, Use of wavelet and fast fourier transforms in pharmacodynamics, J. Pharmacol. Exp. Ther. 321 (2007) 423–430.

[22] K. Odziomek, D. Ushizima, P. Oberbek, K.J. Kurzydłowski, T. Puzyn, M. Haranczyk, Scanning electron microscopy image representativeness: morphological data on nanoparticles, J. Microsc. 265 (2017) 34–50.

[23] T. Wiseman, S. Williston, J.F. Brandts, L.-N. Lin, Rapid measurement of binding constants and heats of binding using a new titration calorimeter, Anal. Biochem. 179 (1989) 131–137.

[24] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.

[25] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.

[26] S. Forli, R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, A.J. Olson, Computational protein–ligand docking and virtual drug screening with the AutoDock suite, Nat. Protoc. 11 (2016) 905.

[27] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, PubChem substance and compound databases, Nucleic Acids Res. 44 (2015) D1202–D1213.

[28] J. Jiménez, S. Doerr, G. Martínez-Rosell, A. Rose, G. De Fabritiis, DeepSite: protein binding site predictor using 3D-convolutional neural networks, Bioinformatics 33 (2017) 3036–3042.

[29] A. Tao, Y. Huang, Y. Shinohara, M.L. Caylor, S. Pashikanti, D. Xu, ezCADD: a rapid 2D/3D visualization-enabled web modeling environment for democratizing computeraided drug design, J. Chem. Inf. Model. 59 (2018) 18–24.

[30] R. Thomsen, M.H. Christensen, MolDock: a new technique for high-accuracy molecular docking, J. Med. Chem. 49 (2006) 3315–3321.

[31] W.P. Feinstein, M. Brylinski, Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets, Journal of cheminformatics 7 (2015) 18.

[32] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, ACS Publications 51 (10) (2011) 2778–2786.

[33] U. Emekli, D. Schneidman-Duhovny, H.J. Wolfson, R. Nussinov, T. Haliloglu, HingeProt: automated prediction of hinges in protein structures, Proteins: Structure, Function, and Bioinformatics 70 (2008) 1219–1227.

[42] W. Zheng, S. Doniach, A comparative study of motor-protein motions by using a simple elastic-network model, Proc. Natl. Acad. Sci. 100 (2003) 13253–13258.
[43] D. Ming, M.E. Wall, Allostery in a coarse-grained model of protein dynamics, Phys. Rev. Lett. 95 (2005), 198103.
[44] J. de Ruyck, G. Brysbaert, R. Blossey, M.F. Lensink, Molecular docking as a popular tool in drug design, an in silico travel, Advances and applications in bioinformatics and chemistry: AABC 9 (2016) 1.
[45] J. Li, Q. Du, C. Sun, An improved box-counting method for image fractal dimension estimation, Pattern Recogn. 42 (2009) 2460–2469.
[46] M. Alia, A. Samsudin, New key exchange protocol based on Mandelbrot and Julia fractal set, International journal of computer science and network security 7 (2007) 302–307.
[47] Y. Fisher, Fractal Image Compression: Theory and Application, Springer Science & Business Media, Place Published, 2012.
[48] B.B. Mandelbrot, Plane DLA is not self-similar; is it a fractal that becomes increasingly compact as it grows? Physica A: Statistical Mechanics and its Applications 191 (1992) 95–107.
[49] B.B. Mandelbrot, Fractal aspects of the iteration of z→ Λz (1-z) for complex Λ and z, Ann. N. Y. Acad. Sci. 357 (1980) 249–259.
[50] D. Saupe, Efficient computation of Julia sets and their fractal dimension, Physica D: Nonlinear Phenomena 28 (1987) 358–370.
[51] J. Soddell, R. Seviour, Using box counting techniques for measuring shape of colonies of filamentous micro-organisms, Complex. Int. 2 (1995) 1–8.
[52] M. Nezadal, O. Zemeskal, M. Buchnicek, The Box-counting: Critical Study, 4th Conference on Prediction, Synergetic and More, The Faculty of Management, Institute of Information Technologies, Faculty of Technology, Tomas Bata University in Zlin, HarFA software, October, 2001 25–26.
[53] S. Preus, K. Kilså, F.-A. Miannay, B. Albinsson, L.M. Wilhelmsson, FRETmatrix: a general methodology for the simulation and analysis of FRET in nucleic acids, Nucleic Acids Res. 41 (2013) e18.
[54] N. Hassan, M.P. Gárate, T. Sandoval, L. Espinoza, Á. Piñeiro, J.M. Ruso, On the selfassembly of a highly selective benzothiazole-based TIM inhibitor in aqueous solution, Langmuir 26 (2010) 16681–16689.

[55] L.R. Barbosa, M.G. Ortore, F. Spinozzi, P. Mariani, S. Bernstorff, R. Itri, The importance of protein-protein interactions on the pH-induced conformational changes of bovine serum albumin: a small-angle X-ray scattering study, Biophys. J. 98 (2010) 147–157.

[56] J.M. Kollman, L. Pandi, M.R. Sawaya, M. Riley, R.F. Doolittle, Crystal structure of human fibrinogen, Biochemistry 48 (2009) 3877–3886.

[57] M. González-Durruthy, G. Scanavachi, R. Rial, Z. Liu, M.N.D.S. Cordeiro, R. Itri, J.M. Ruso, Structural and energetic evolution of fibrinogen toward to the betablocker interactions, Int. J. Biol. Macromol. 137 (2019) 405–419.

[58] R. Itri, L.Q. Amaral, Distance distribution function of sodium dodecyl-sulfate micelles by X-ray-scattering, J Phys Chem-Us 95 (1991) 423–427.

[59] D. Franke, M.V. Petoukhov, P.V. Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, D.I. Svergun, ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions, J. Appl. Crystallogr. 50 (2017) 1212–1225.

[60] N. Sturm, J.r.m. Desaphy, R.J. Quinn, D. Rognan, E. Kellenberger, Structural insights into the molecular basis of the ligand promiscuity, J. Chem. Inf. Model. 52 (2012) 2410–2421.

[61] A.M. Afzal, H.Y. Mussa, R.E. Turner, A. Bender, R.C. Glen, A multi-label approach to target prediction taking ligand promiscuity into account, Journal of cheminformatics 7 (2015) 24.

[62] A.M. Afzal, H.Y. Mussa, R.E. Turner, A. Bender, R.C. Glen, A multi-label approach to target prediction taking ligand promiscuity into account, Journal of cheminformatics 7 (2015) 24.

[63] P. Källblad, R.L. Mancera, N.P. Todorov, Assessment of multiple binding modes in ligand− protein docking, J. Med. Chem. 47 (2004) 3334–3337.

[64] C. Chennubhotla, A. Rader, L.-W. Yang, I. Bahar, Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies, Phys. Biol. 2 (2005) S173.

[65] B.H. Lee, S. Seo, M.H. Kim, Y. Kim, S. Jo, M.-k. Choi, H. Lee, J.B. Choi, M.K. Kim, Normal mode-guided transition pathway generation in proteins, PLoS One 12 (2017) e0185658.

[66] T. Oliwa, Y. Shen, cNMA: a framework of encounter complex-based normal mode analysis to model conformational changes in protein interactions, Bioinformatics 31 (2015) i151–i160.
[67] O. Keskin, S.R. Durell, I. Bahar, R.L. Jernigan, D.G. Covell, Relating molecular flexibility to function: a case study of tubulin, Biophys. J. 83 (2002) 663–680.
[68] R.C. Gonzalez, R.E. Woods, Digital image processing [M], Publishing house of, Electron. Ind. 141 (2002).
[69] A. Malliani, M. Pagani, F. Lombardi, S. Cerutti, Cardiovascular neural regulation explored in the frequency domain, Circulation 84 (1991) 482–492.
[70] J.L. Semmlow, Biosignal and Medical Image Processing, CRC press, Place Published, 2008. [65] G. Weng, U.S. Bhalla, R. Iyengar, Complexity in biological signaling systems, Science 284 (1999) 92–96.
[71] H. Li, S. Chen, H. Zhao, Fractal structure and conformational entropy of protein chain, Int. J. Biol. Macromol. 12 (1990) 374–378.
[72] A.D. Koromyslova, A.O. Chugunov, R.G. Efremov, Deciphering fine molecular details of proteins' structure and function with a Protein Surface Topography (PST) method, J. Chem. Inf. Model. 54 (2014) 1189–1199.
[73] A. Malliani, M. Pagani, F. Lombardi, S. Cerutti, Cardiovascular neural regulation explored in the frequency domain, Circulation 84 (1991) 482–492.
[74] J.L. Semmlow, Biosignal and Medical Image Processing, CRC press, Place Published, 2008. [65] G. Weng, U.S. Bhalla, R. Iyengar, Complexity in biological signaling systems, Science 284 (1999) 92–96.
[75] H. Li, S. Chen, H. Zhao, Fractal structure and conformational entropy of protein chain, Int. J. Biol. Macromol. 12 (1990) 374–378.
[76] A.D. Koromyslova, A.O. Chugunov, R.G. Efremov, Deciphering fine molecular details of proteins' structure and function with a Protein Surface Topography (PST) method, J. Chem. Inf. Model. 54 (2014) 1189–1199.
[77] R. Itri, L.Q. Amaral, Distance Distribution Function of Sodium Dodecyl-Sulfate Micelles by X-Ray-Scattering, J Phys Chem-Us, 95 (1991) 423-427.
[78] G. Rabbani, E. Ahmad, N. Zaidi, S. Fatima, R.H. Khan, pH-Induced molten globule state of Rhizopus niveus lipase is more resistant against thermal and chemical denaturation than its native state, Cell Biochem. Biophys. 62 (2012) 487–499.
[79] A. Guinier, G. Fournet, Small-Angle Scattering of X-rays, Wiley, Place Published, 1955.

Supporting Information

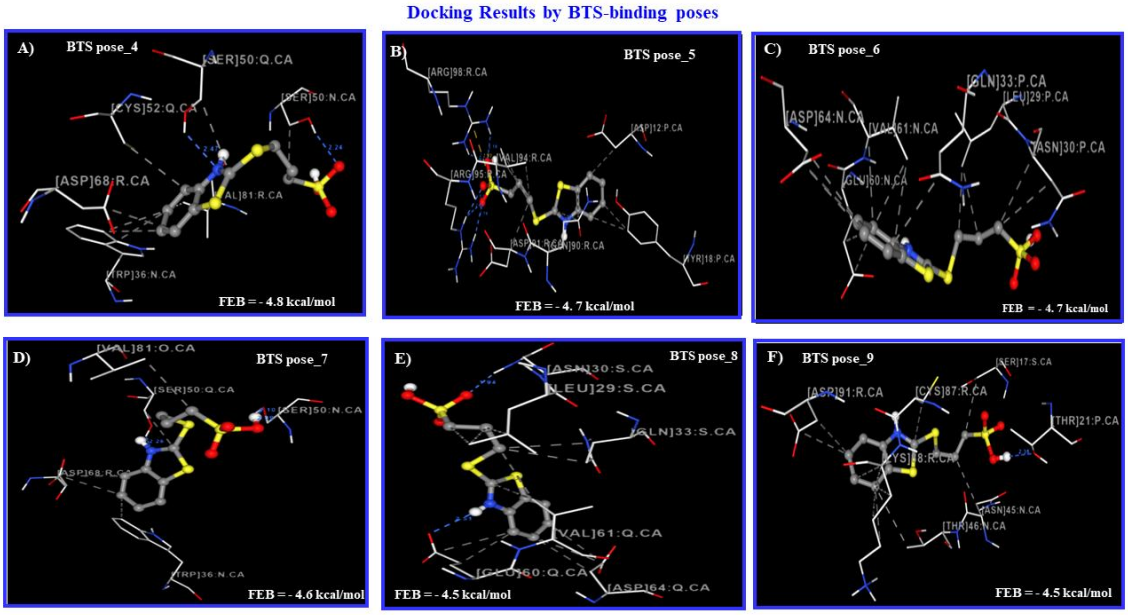


Figure S1. Results of molecular docking simulation between BTS_binding poses and fibrinogen E-region. **A)** BTS_*pose 4*, **B)** BTS_*pose 5*, **C)** BTS_*pose 6*, **D)** BTS_*pose 7*, **E)** BTS_*pose 8*, and **F)** BTS_*pose 9* obtained from the docking complexes BTS-fibrinogen depicted with their corresponding FEB values in kcal/mol. Herein gray dotted-lines between the different BTS-poses and fibrinogen residue interacting-atoms represent Van der Waals hydrophobic interactions and the blue dotted-lines represent H-bond contacts.

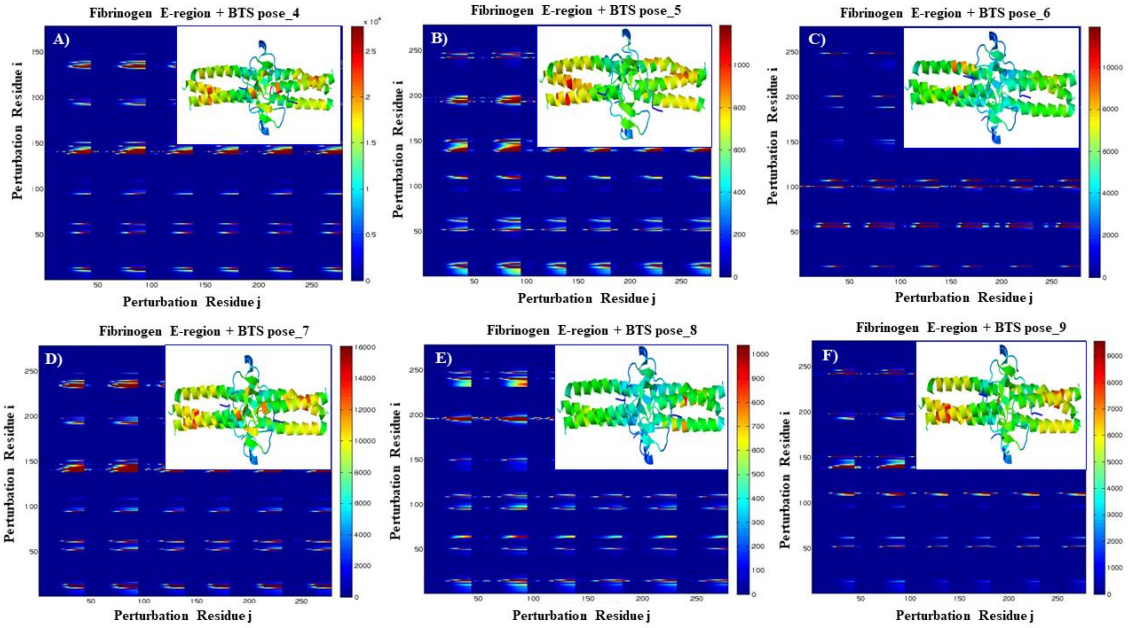


Figure S2. Perturbation response analysis for BTS-binding poses. **A-F**) Individual LPRS-maps obtained from the BTS-fibrinogen docking complexes-based conformation analysis depicted on the top right rectangle of the LPRS-maps with the corresponding intensity bar color (on the right). For this purpose, dark blue to red regions correspond to low to high strength of perturbation response respectively for 278 residues from the fibrinogen E-region like by C(α)-atoms. All the LPRS-maps were established in range of the low frequency normal modes = 2.

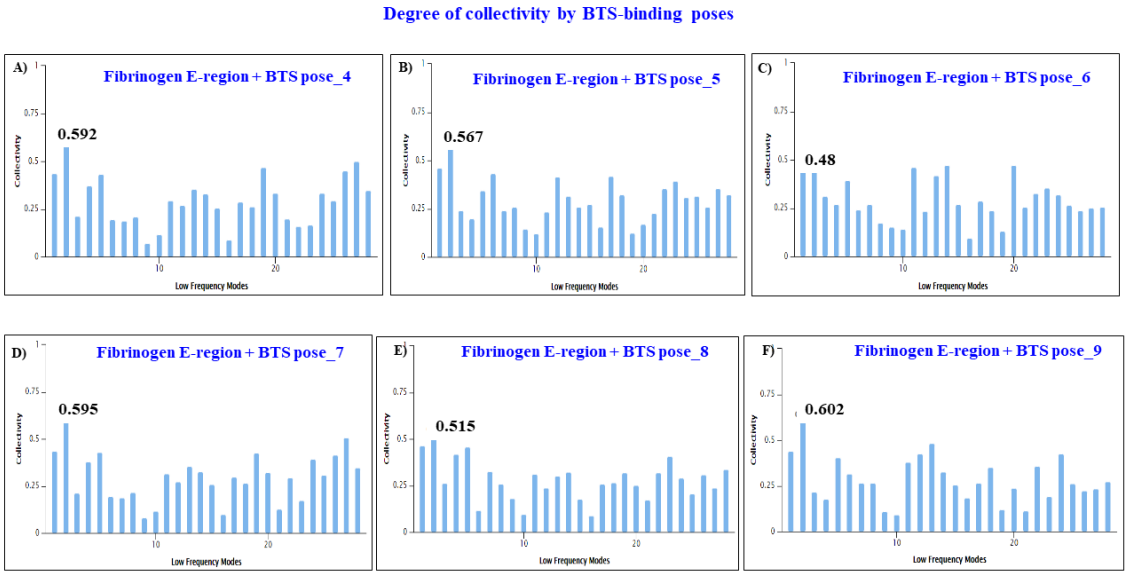


Figure S3. Graphical representation of the degree of collectivity *K*(*k*) *vs.* low frequency normal modes (mode 2) of the fibrinogen E-region interacting with the different BTS-

binding poses like: **A)** Fibrinogen + BTS_*pose 4*, **B)** Fibrinogen + BTS_*pose 5*, **C)** Fibrinogen + BTS_*pose 6*, **D)** Fibrinogen + BTS_*pose 7*, **E)** Fibrinogen + BTS_*pose 8*, **F)** Fibrinogen + BTS_*pose 9* (0.48).

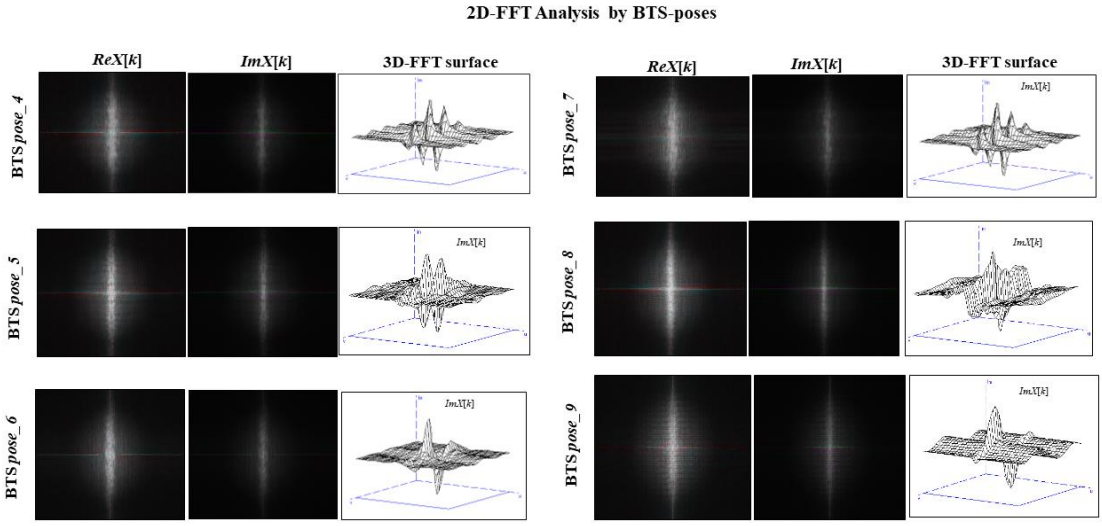


Figure S4. Graphical representation of the 2D-FFT profiles like: FT phases from the frequency domain ($X[k]$), $ReX[k]$, $ImX[k]$, and the 3D-FT surface graph like $F(u,v) = X[k]$; with $ImX[k] \neq 0$ and $ReX[k] = 0$ obtained from each LPRS-map of BTS-crystallographic conformational binding mode. From top to botton and left to right: Fibrinogen + BTS_*pose 4*; Fibrinogen + BTS_*pose 5*; Fibrinogen + BTS_*pose 6*; Fibrinogen + BTS_*pose 7*; Fibrinogen + BTS_*pose 8*; Fibrinogen + BTS_*pose 9*. For all the 3D-FT surface graph the coordinates axes (u, v) are included into the spatial frequency and represent the real and imaginary dimension, respectively.

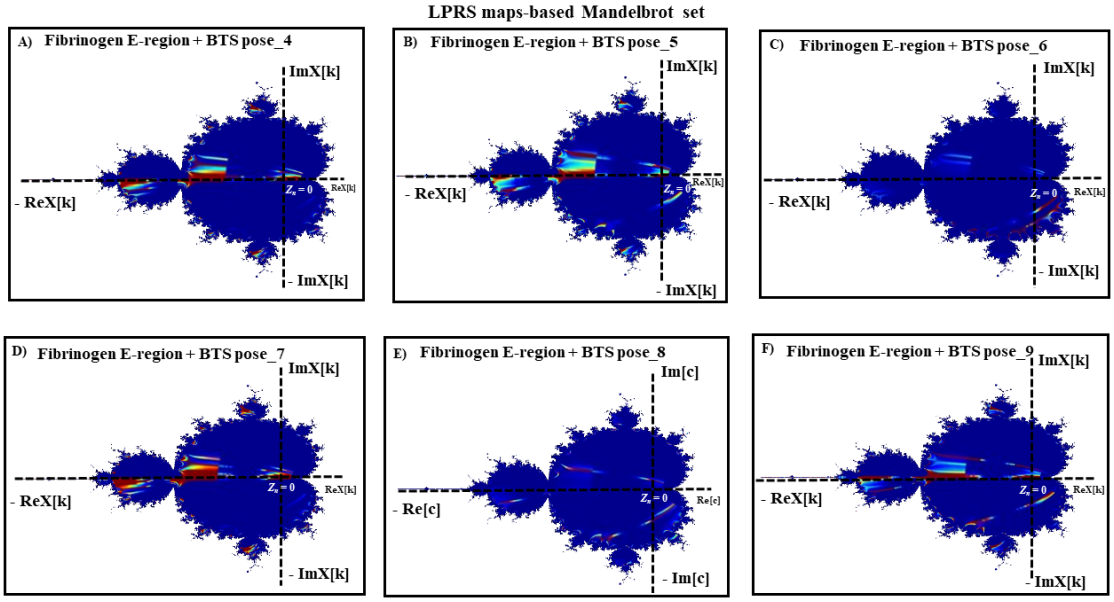


Figure S5. LPRS maps data visualization-like Mandelbrot set fractal approach with cardioid image modeling. **A)** Fibrinogen + BTS_*pose 4*, **B)** Fibrinogen + BTS_*pose 5*, **C)** Fibrinogen + BTS_*pose 6*, **D)** Fibrinogen + BTS_*pose 7*, **E)** Fibrinogen + BTS_*pose 8*, **F)** Fibrinogen + BTS_*pose 9*. The Mandelbrot set parameters were fixed in $ReX[k]_{max}$ = 1, $ReX[k]_{min}$ = -2, $ImX[k]_{max}$ =1, $ImX[k]_{min}$ = -1 for modeling the main cardioids.

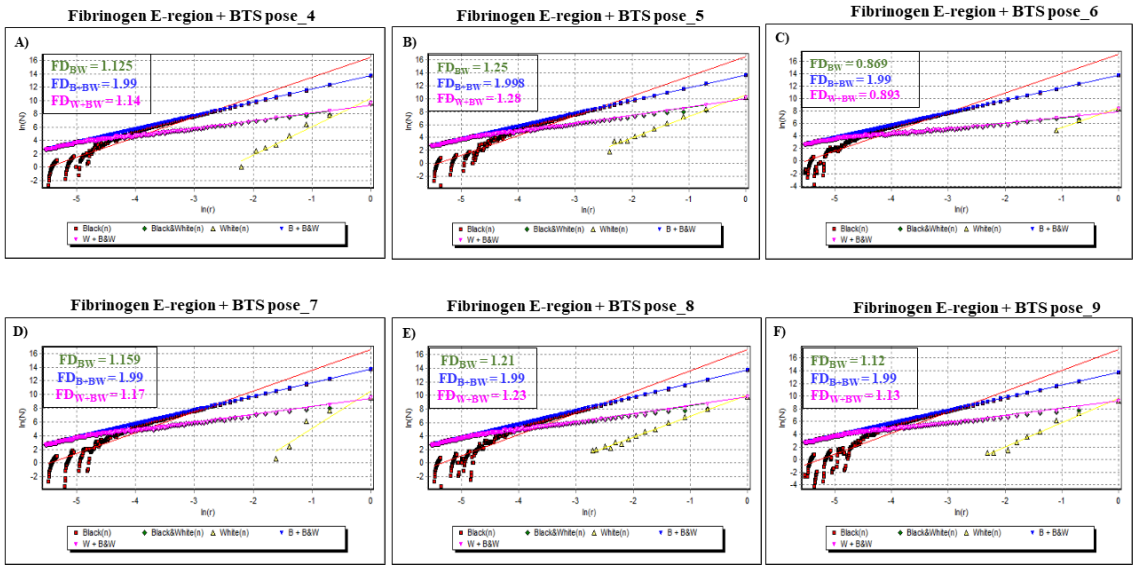


Figure S6. Fractal spectrum-based on box-counting algorithm applied to obtain the slopes like fractal dimensions (*FD*: D_{BW}, D_{BBW}, and D_{WBW}) of the linear regression yields from

binary black/white binary LPRS maps image-processing for the BTS-conformational ligand binding-modes like **A)** Fibrinogen + BTS_*pose 4*, B**)** Fibrinogen + BTS_*pose 5*, C**)** Fibrinogen + BTS_*pose 6*, D**)** Fibrinogen + BTS_*pose 7*, E**)** Fibrinogen + BTS_*pose 8*, F**)** Fibrinogen + BTS_*pose 9*.

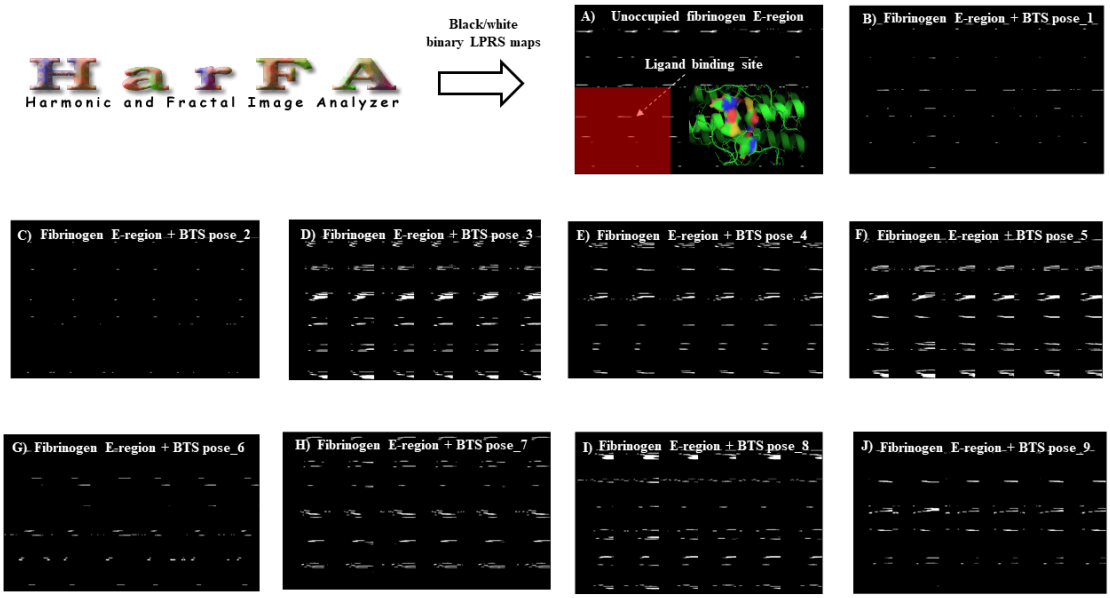


Figure S7. Black/white binary LPRS map imagen-processing used to perform the fractal box-counting analysis in this study. **A)** Unoccupied fibrinogen E-region, **B)** Fibrinogen + BTS_*pose 1*, **C)** Fibrinogen + BTS_*pose 2*, **D)** Fibrinogen + BTS_*pose 3*, **E)** Fibrinogen + BTS_*pose 4*, **F)** Fibrinogen + BTS_*pose 5*, **G)** Fibrinogen + BTS_*pose 6*, **H)** Fibrinogen + BTS_*pose 7*, **I)** Fibrinogen + BTS_*pose 8*, **J)** Fibrinogen + BTS_*pose 9*.

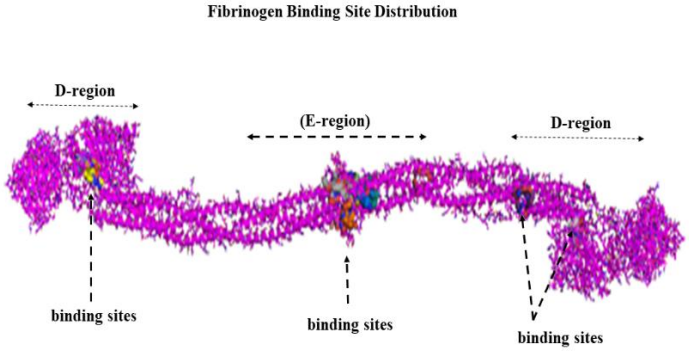


Figure S8. Representation of general crystallographic structure of fibrinogen protein by relevant regions formed by two quasi-symmetric C-terminal portions (D-regions) and the central fragment (E-region thrombin binding-domain) highlighting the binding sites formed by small cluster of hydrophobic cavities represented as colored van der Waals surfaces heterogeneously distributed across the overall structure.

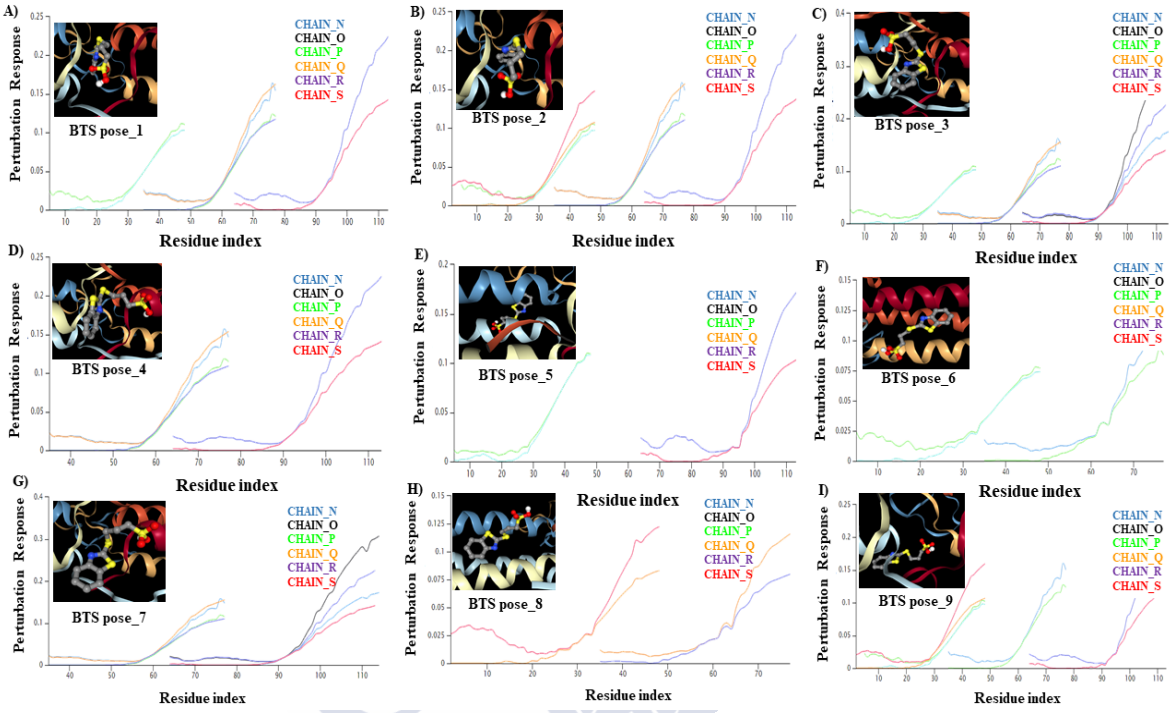


Figure S9. Graphical representation of the different perturbation responses induced by different BTS-orientation binding-modes or BTS crystallographic poses over fibrinogen E-region chains.

Chapter 5.

Michael González-Durruthy, Riccardo Concu, Laura F. Osmari Vendrame, Ivana Zanella, Juan M. Ruso, M. Natália D.S. Cordeiro. Targeting beta-blocker drug-drug interactions with fibrinogen blood plasma protein: A computational and experimental study, Molecules, 25, 5425, 2020. doi:10.3390/molecules25225425.

Journal Impact Factor (JIF) (2020): 4.411

CiteScore (2020): 4.1

Author Contribution M.G.-D.; Conceptualization, methodology-based on molecular docking coupled DFT and computational polypharmacology simulations, building perturbation response scanning maps, and writing—original draft preparation. *Corresponding author

Journal authorization:

Chapter 5.

Computational study on beta-blocker drug-drug interactions on fibrinogen blood plasma protein

Abstract:

In this work, one of the most prevalent polypharmacology drug-drug interaction which occurs between two widely used beta-blocker drugs (i.e., ligand-ligand binding event for acebutolol plus propranolol with the most abundant blood plasma fibrinogen protein) was evaluated. Towards that end, molecular docking and Density Functional Theory (DFT) calculations were used as complementary tools. A fibrinogen crystallographic validation for the three best ranked binding-sites shows 100% of conformationally favored residues with total absence of restricted flexibility. From those three sites, results on both the binding-site druggability and ligand transport analysis-based free energy trajectories pointed out the most preferred biophysical environment site for drug-drug interactions. Furthermore, the total affinity for the stabilization of the drug–drug complexes was mostly influenced by steric energy contributions, based mainly on multiple hydrophobic contacts with critical residues (THR22: P and SER50: Q) in such best-ranked site. Additionally, the DFT calculations revealed that the beta-blocker drug-drug complexes have a spontaneous thermodynamic stabilization following the same affinity order obtained in the docking simulations, without covalent-bond formation between both interacting beta-blockers in the best-ranked site.

5.1 Introduction

Computational polypharmacology drug-drug interaction (DDI) studies are essential in rational drug-design and development [1–7]. Yet the exhaustive exploration of drug-drug side effects by experimental techniques remains a great challenge for the pharmaceutical industry owing to the involved costs and time. For example, beta-blocker DDI events constitute one of the most prevalent side effects associated with complications in several pharmaceutical therapies such as those employed to treat cardiovascular diseases. From the pharmacodynamics point of view, the DDI events refer to the simultaneous interactions of more than one drug (belonging to the same or different pharmacological groups) in the same molecular binding site of a given molecular target [2,8–10]. Polypharmacology DDI facets can affect millions of patients each year and apparently,

the most feasible approach to address them is to use computational methods [4–6]. In fact, advanced computational methods such as machine learning, molecular dynamics, pharmacophore modeling, and molecular docking techniques have become powerful in silico tools at the initial stages of rational drug design [11–17]. Currently, the main limitation of these methods is to cope with the mechanistic interpretation of DDI phenomena at the atomic level [11–14]. Indeed, computational polypharmacology involves the characterization of DDI across multiple scales and organization levels by integrating drug specific binding properties [7]. Nevertheless, advanced molecular docking simulations combined with Density Functional Theory (DFT) calculations can be applied efficiently to model DDI events on the relevant binding sites of a given protein with atomic precision. Specifically, in DDI studies, molecular docking simulations allow one to identify the stable drug–drug complexes by firstly predicting the feasible binding-sites of a target protein [13,17], followed by establishing from those, the best-ranked drug–drug binding configurations as well as their orientations within the best-ranked binding-sites along with the assessment of their binding affinity [11,18–20]. During the ligand-binding processes, it is common to find propensity of the ligands to interact with multiples binding sites within the same protein. The greater or lesser specificity of a given ligand to a particular binding site relies on the affinity at the biophysical environment where the interaction takes place, and the latter can be assessed by the druggability degree (Dg) of the binding site [18–20]. Dg represents a quantitative estimation of the maximum intrinsic affinity of a given binding site of the protein to bind one or more drugs (ligands). Notice, however, that several pockets may have the ability to bind more than one drug enabling the occurrence of pharmacodynamics DDI, when partial or total overlapping of the drugs occurs at the same biophysical environment of a given binding site of the protein. Despite the existence of great advances in computational pharmacology, how DDI phenomena can modulate the adverse effects at the molecular and atomistic level are still ignored. Previous experimental studies show that beta-blocker drug interactions can induce multiple adverse effects [21–23]. Particularly, when the beta-blockers acebutolol (A) and propranolol (P) are simultaneously administered in the bloodstream, they can induce different clotting perturbations due to interactions with proteins [24,25]. The most abundant protein circulating in the bloodstream plasma is fibrinogen, a protein with a molecular weight of ~340 kDa depending on Aα and Bβ chains and the γ-γ and α-α crosslinking chains content [26]. From a structural and functional point of view, fibrinogen is a soluble glycoprotein with a central nodule (E-region) comprising several

cavities and tunnels, which facilitate the ions, water solvent, and ligand transport, and in turn, its E-region has the maximum responsibility for the fibrin polymerization during the clotting process [27–31]. Therefore, the binding site localized in the most central part of the E-region of fibrinogen protein exhibits the greatest propensity for the occurrence of pharmacodynamics DDI leading to most of the side-effects among A and P beta-blockers. All the theoretical evidences were well-supported by a combination of computational modeling approaches based on molecular docking and DFT calculations allowing to study the polypharmacology drug–drug interactions between the beta-blockers A plus P with the fibrinogen that could induce undesired effects in the coagulation processes where the fibrinogen thrombin binding-domain (central E-region) is critically involved.

5.2 Results and discussion

One of the most important steps in computational rational drug-design is the precise prediction and characterization of the biological target structure [32]. A deep knowledge of its three-dimensional structure, flexibility properties, and residues composition of its catalytic binding site is crucial to understand the phenomena of ligand–protein interactions with relevance from the pharmacological point of view [33–35]. As referred before, during the ligand-binding processes in a given protein target, it is very common to find propensity of the ligand to interact with multiples catalytic binding sites (tiny cavities and tunnels) in the same protein [11–17]. Indeed, the greater or lesser specificity and selectivity of the ligand will be decided by the affinity at these biophysical environments where the interactions take place. Particularly, the DDI phenomena are crucial to prevent side effects and ensure the effectiveness and safety of drugs [1–7]. In this work, the study of polypharmacology DDI between A and P beta-blockers is tackled considering the most relevant fibrinogen E-region binding sites. Towards such purpose, we performed a pocket ranking of fibrinogen E-region binding sites by assessing their highest druggability Dg [36], as defined in the **Material and Methods** section (see **equation (4)**). Notice that the higher Dg values are the maximum intrinsic affinity of the predicted binding sites of fibrinogen E-region with the ability to simultaneously bind the acebutolol and propranolol beta-blockers [36].

5.2.1 Prediction of the Binding Sites for Fibrinogen

Firstly, we present the predictions of the best-ranked catalytic binding sites of receptor fibrinogen obtained using the machine learning technique 3D-DCNN [32]. The 3D-

DCNN algorithm performs a Delaunay triangulation with weighted points based on the computational geometric concept of the "alpha sphere" center [32,36,37]. An alpha sphere is a sphere which contains no internal atom and in turn, links four atoms on its border. In other words, the four atoms are at an equal distance to the alpha sphere center [32,36,37]. Then, the algorithm can detect plausible fibrinogen binding pockets by summing up all alpha sphere centers and later delimiting the access to cavities by scanning the van der Waals surfaces in the fibrinogen E-region. The procedure includes all the depth concave regions of fibrinogen that are directly associated with a high probability to ligand binding (higher Dg values) and exclude all the convex parts associated to low druggability [18–20]. The 3D-DCNN results showed that from a total of ten predicted binding sites in the fibrinogen E-region, the three best-ranked sites are the following ones depicted in Figure 1, namely: site 1 (Dg = 0.81) > site 2 (Dg = 0.54) > site 3 (Dg = 0.39)

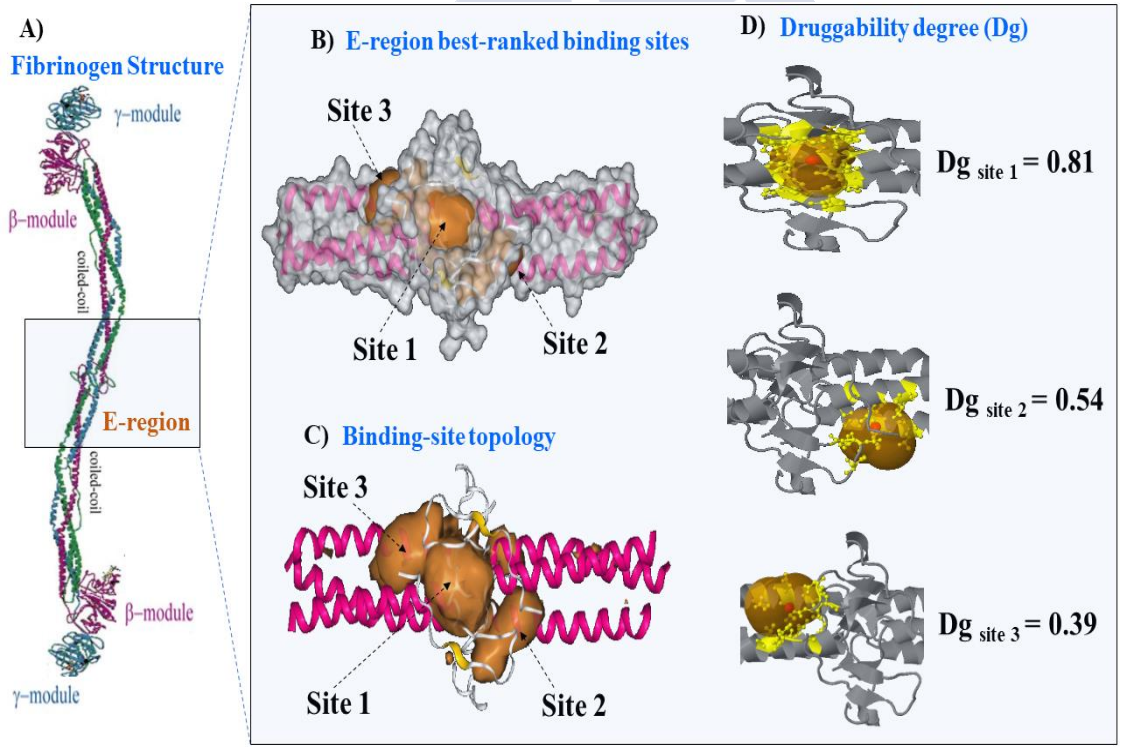


Figure 1. A) Details on the molecular structure of fibrinogen protein with the relevant regions, such as the two C-terminal portions of the carbohydrate-linking coiled-coil like chains (γ-module and β-module) and the funnel hydrophobic cavity thrombin binding-domain (central E-region), the latter highlighted with the light-blue rectangle. **B)** 3D-DCNN prediction of the three most relevant fibrinogen E-region binding-pockets (site 1, site 2 and site 3) depicted as van der Waals surfaces. **C)** Representation of the binding-

site topology linked to catalytic active sites, depicted as volumetric orange regions. **D)** Prediction of the corresponding D_g for the best three catalytic binding sites of the fibrinogen E-region, depicted as orange transparent hull surfaces. Therein, the *alpha sphere* centers for predictions are shown as small red points within each site and surrounding target residues (yellow colored). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

Let us now proceed to the 3D X-ray crystallographic structure validation of the fibrinogen binding sites. Figure 2 shows the Ramachandran plots for its three best-ranked binding sites. These plots contain a 2D-projection on the plane obtained from the 3D-crystallographic binding sites, pertaining to all their possible residues conformations according to the dihedral angles (Psi (ψ) vs. Phi (ϕ)) around the peptide bond of such residues. The allowed torsion values for the ψ vs. ϕ dihedral angles, placed within the Ramachandran colored purple contour, are identified as conformationally favored residues, otherwise, these are considered as sterically disallowed residues [38]. This validation has paramount importance in order to avoid the presence of potential false positives in the docking results coming from the target residues forming the E-region binding sites.

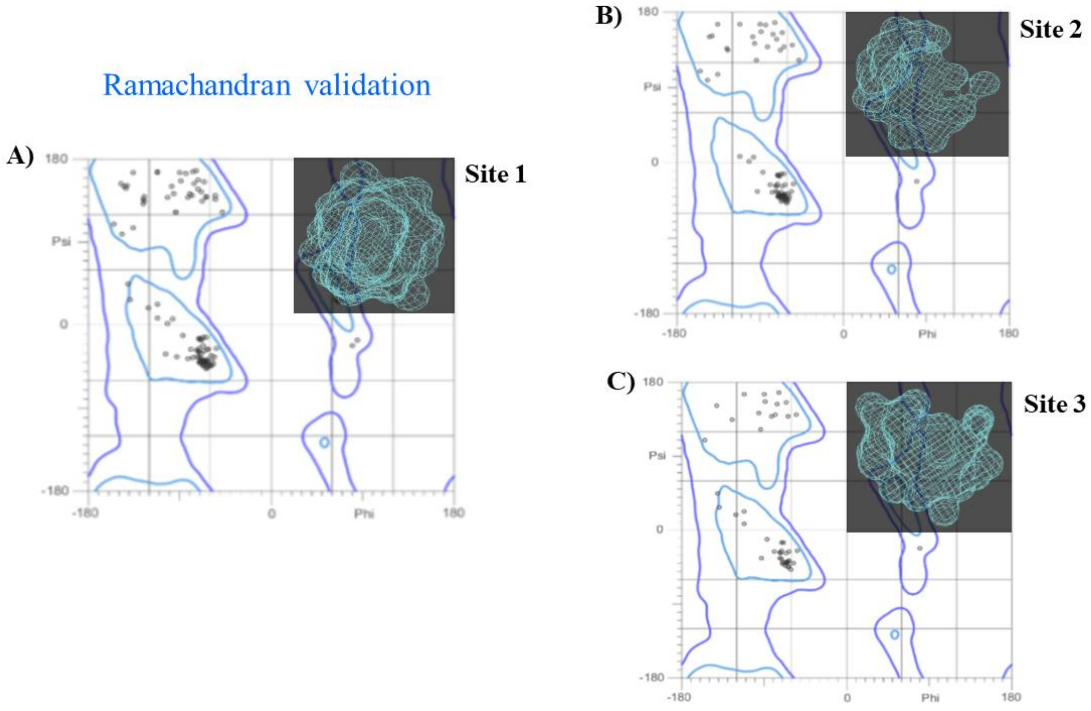


Figure 2. Ramachandran diagrams obtained for the best-ranked fibrinogen binding sites: **A)** site 1, **B)** site 2 and **C)** site 3. For visualization purposes, these are also 3D-

displayed in the upper right corner (light blue) of each panel. All the possible combinations of torsion dihedral angles Psi (ψ) *vs.* Phi (φ) are shown. Note the total absence of Ramachandran outlier residues for the three sites analyzed, usually located outside the Ramachandran colored purple contour, if any. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

As can be seen in Figure 2, our results show that 100 % of the residues belonging to the three validated active sites can be categorized like conformationally favored residues. Moreover, the complete lack of restricted flexibility of such residues will clearly avoid carrying out false positive docking studies of the beta-blocker DDI. Here, it is important to notice that, in most docking studies, the Ramachandran validation is absent or in some cases inappropriately applied over all the protein. In this work, the Ramachandran X-ray crystallographic validation was done for only the clusters of residues composing the relevant fibrinogen E-region binding sites, then allowing to obtain a more adjusted evaluation about the conformational integrity of the target-residues potentially involved in the docking interactions.

5.2.2 Identification of tunnels for the fibrinogen binding sites.

Fibrinogen E-region binding-sites are highly complex systems containing a large and variety number of tunnels, clefts, grooves, protrusions, and empty spaces in the interior of the tangled funnel hydrophobic cavity formed by the γ-γ and α-α crosslinking from Aα and Bβ chains. These chains are located in the two outer sides of the binding sites for thrombin, which allow the supercoiling of protofibrils in the fibrinogen molecule during critical stages of the blood coagulation process [26–29]. Under the presence of ligands such as the A and P here studied, critical binding residues can be affected by modifications in the symmetry architecture of the Bβ-γ/Bβ-γ dimeric domain of E-region. In fact, it is well known that many β-adrenoreceptor blocking agents can affect the parameters of the blood coagulation such as fibrinogen concentration and activity when co-administered intravenously [24]. Let us now then focus our attention on the structural prediction and characterization of the tiny cavities (tunnels) composing the catalytic residues with the highest relevance-based druggability. From the structural point of view, the tunnels are defined as entangled and small sub-cavities occupying the internal volume of the aforementioned binding sites, which selectively delimit the ligand transport to them, and usually present a narrower access determining the catalytic properties of the

binding sites of the fibrinogen E-region. The specific functions of the tunnels include: (i) exchange of solutes, ions, and water between the tiny cavities; (ii) leading the transport of reaction intermediates between two distinct active sites; and (iii) maintaining the conditions of conformational and geometric complementarity for binding physiological substrates. Moreover, the residues forming the tunnel are considered as promising hotspots with high ligand-binding propensity [36]. Thus, taking into account the complexity of the tunnel hydrophobic cavity of the fibrinogen E-region [26–29], evaluating the catalytic function of fibrinogen residues on its detected tunnels is considered to be more informative than that of traditional predicted topological cavities. According to this and by resorting to Caver software [36], we identified at least two tunnels geometrically interconnected for each predicted catalytic binding-site (sites 1–3), sharing catalytic binding residues (see **Figure 3**).

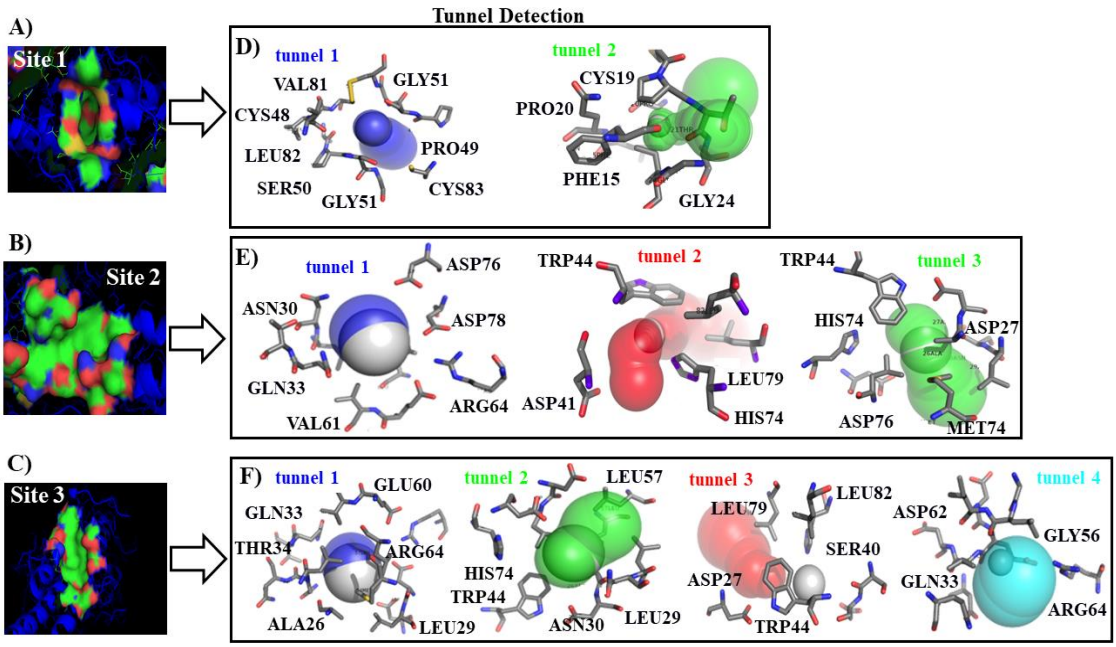


Figure 3. On the left, van der Waals surface representation of the three best-ranked fibrinogen binding-sites, namely: (A) site 1, (B) site 2, and (C) site 3. Therein, red and blue regions represent acid and basic residues, respectively. On the right, (D–F) representation of the different tiny cavities such as tunnels detected for each predicted fibrinogen binding site with the corresponding surrounding target-residues. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

It is important to highlight here that binding sites with high druggability comprise target-residues displaying higher tolerance for deep environments such as: CYS, ALA, ILE,

LEU, MET, PHE, VAL, and TRP. The latter in turn directly influence the ligand binding properties of the established protein–ligand complexes (A + P/fibrinogen)—i.e., their specificity, enantioselectivity, inhibition constants, and thermodynamic stability [39,40]. As mentioned before, we also identified the "worst fibrinogen binding site" to be used as reference control denoting the absence of beta-blocker drug–drug interactions. For this instance, a small region formed by two different tunnels located in opposite crystallographic planes of the fibrinogen E-region was identified and showed to have the lowest value of druggability degree (Dg = 0.08). More details concerning the worst fibrinogen binding site can be found in **Figure S1** of the Supplementary Material.

5.2.3. Calculation of Energetic Contributions for Binding Affinity

Moving on to the docking experiments, now we will focus our attention on the study of potential DDI phenomena between A and P beta-blockers with the receptor fibrinogen E-region. Since several tunnels from the three binding sites were detected in the fibrinogen E-region (**Figure 3**), we established three pocket ranking prioritization criteria. The first one is the mutual overlap criterion (MOc) [37], which considers a given binding site likely relevant for DDI, if partial or total ligand-overlapping exists at the same crystallographic plane or biophysical environment in that site. Additionally, it considers the site that is likely relevant, if at least > 50% of the ligand atoms are within 3 Å of at least one alpha sphere. The second criterion pertains to the ligand-docking complexes with spontaneous thermodynamic process ($\Delta G_{bind} < 0$ kcal/mol). With regard to this criterion, as already referred to, the docking simulation results are categorized as energetically unfavorable when the Gibbs free energy of the formed complexes is $\Delta G_{bind} \geq 0$ kcal/mol, pointing either to extremely low or complete absence of binding affinity, otherwise they are categorized as having medium to high docking affinity [41,42]. Finally, the third criterion relates to the high depth of the target-residues of relevant binding sites. The depth is defined as the distance of any atom in a given residue to its closest water molecule from virtual bulk solvent [39,40]. In this instance, it is important to note that the cited crystallographic water molecules were removed from the binding sites before starting the molecular docking simulations during the step of receptor preparation (see Materials and Methods section). Several factors can influence computational polypharmacology DDI studies. Amongst others are the protein's crystallographic quality, its flexibility properties, the pharmacological relevance of the catalytic residues involved in the polypharmacology interactions, the physico-chemical properties of the interacting

ligands, and the presence of different conformations within the binding sites. Particularly, the thermodynamic factors are crucial to tackle the polypharmacology DDI mechanisms. To address the latter, we carried out a ligand transport analysis for the obtained DDI complexes between the beta-blockers A and P with the three best-ranked binding-sites. The LTA-approach executes an iterative docking algorithm which simultaneously scans the binding-energy trajectory vs. tunnel radius under spatially restrained conditions, considering the same biophysical environment for both beta-blockers to every slice of the binding sites (see **Figure 4**). In general terms, the obtained beta-blocker binding energy trajectories fit with a spontaneous thermodynamic process (ΔGbind < 0 kcal/mol), therefore suggesting the formation of stable beta-blocker drug–drug complexes in the three predicted fibrinogen binding sites. According to this, the modeling results on the throughput tunnel parameter did not show significant differences between the three best-ranked binding sites. In fact, the estimated T values from the predicted tunnels were high and very close to each other (i.e., T[site 1-tunnel 2] = 0.90, T[site 2-tunnel 1] = 0.95, and T[site 3-tunnel 1] = 0.96), indicating that the three binding sites display a similar performance for the transport of both beta-blockers. Concerning this, note that T values close to 1 reflect a high probability of these tunnels for the simultaneous transport of beta-blocker ligands increasing the probability of DDI. However, considering the obtained ΔGbind values from the LTA-analysis, one can clearly distinguish that the DDI phenomenon hierarchically follows the order: site 1 > site 2 > site 3. In the case of the trajectories through the binding site 1, the obtained binding energy profiles for both beta-blockers suggest the formation of a DDI complex with a high stability with a range of ΔGbind energy starting from −5.8 to −7.40 kcal/mol. Additionally, the interaction affinity soared significantly in the binding site 1 around the tunnel radius (≈2.83 Å) until its critical value (≈2.4 Å) for both beta-blockers, which fits with the maximum of drug–drug coupling interaction near to the narrowest part of the binding site 1. After this, the drug–drug trajectories became abruptly independent for both beta-blockers. Interestingly, beta-blocker A continued interacting at site 1 and increased its binding affinity from −7.40 kcal/mol until a maximum value of ΔG_{bind} = −7.70 kcal/mol. In contrast, P showed several peaks of significant loss of affinity in the energy range of −7.40 and −6.5 kcal/mol in that site. According to the binding trajectory of P, several factors are likely to explain its behavior, such as rotational and repulsive forces linked to different binding modes, as well as clashes of P with the residues surrounding tunnels 1 and 2 of site 1. In fact, these likely induce a relative loss of the P affinity under the influence of A at the same

biophysical environment. Recent experimental evidence using small-angle X-ray scattering (SAXS) also suggests that P significantly affects the fibrinogen protein more than A. However, we firmly believe that A induces a reversible loss of P affinity [43], which can be verified in the range of 7–9 Å for the isolated trajectory of P (see **Figure 4D**).

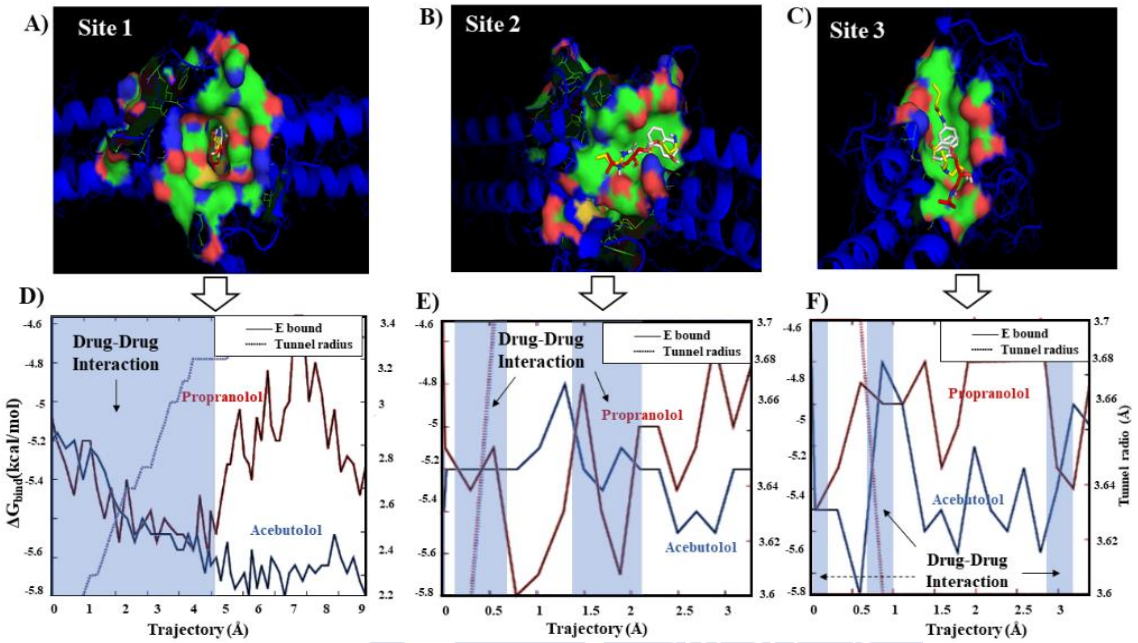


Figure 4. On the top, van der Walls surface representation for the best drug–drug interaction (DDI) systems forming stable docking complexes with the fibrinogen E-region binding sites. (A) A plus P at site 1; (B) A plus P at site 2; and C) A plus P at site 3. On the bottom, panels (D-F) show the individual Gibbs free energy profiles (ΔGbind) from both beta-blockers plotted as a function of the trajectory and the fibrinogen tunnel radius across the predicted catalytic sites. Therein, the corresponding energy profile of A and P is depicted in red and blue solid lines, respectively. The transparent blue rectangle with the black arrow in each panel indicates the occurrence of DDI events (i.e., A and P interacting at the same biophysical environment in sites 1, 2, and 3). For comparison purposes, please refer to **Figure S2**, where the LTA results for the worst fibrinogen binding site are shown. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

Instead, the obtained trajectories for A and P in sites 2 and 3 exhibited a non-continuous DDI profile (i.e., regions where the binding trajectories are independent or non-associated) with an average affinity around −5.22 kcal/mol. In both these sites, a similar behavior happened in the first steps of the beta-blocker transport around tunnel radius

≈3.5 Å. In addition, one can detect the absence of DDI trajectories between 0.5–1.4 Å and 0.2–0.5 Å for sites 2 and 3, respectively. In site 2, the absence of DDI trajectories may be associated with the reduction in the A affinity (ΔGbind ranges from −5.2 to −4.9 kcal/mol) and, in site 3, with the reduction in P affinity (ΔGbind ranges from −5.5 to −4.9 kcal/mol). So, this suggests that these differences between the beta-blocker trajectories could be attributed to their transport through different tunnels, and the depth of the surrounding target-residues composing the tunnels of both sites. From the 2 Å trajectory for both beta-blockers in the binding sites 2 and 3, one can see the absence of DDI in a large part of the transport of these beta-blockers associated to a loss of P affinity, may be due to the presence of A interacting simultaneously with the same target-residues (vide infra), clearly affecting the thermodynamic stabilization of P interactions in such sites. It is important to note that in the binding sites 2 and 3, the apparent similarity of the beta-blocker drug–drug trajectories could be attributed to a quasi-symmetrical architecture of these sites with respect to the main center of tunnel-hydrophobic cavity (site 1) in the fibrinogen E-region [26]. In addition, we checked the mutual overlap criterion by performing a control of simulation experiment based on the LTA approach for the worst fibrinogen binding site (Dg = 0.08) [36,37]. In doing so, it was corroborated that the beta-blocker trajectories for A and P do not intercept during the passage through this site (see **Figure S2**). This result is of great theoretical relevance because it serves as a reference to identify the total absence of DDI patterns when compared to the interactions in sites 1, 2, and 3. Following this idea, to verify the presence or absence of DDI for beta-blockers A and P at the same biophysical environment in the best-ranked fibrinogen binding sites, the corresponding 3D-ligand plot diagrams [44,45] were determined and are shown in **Figure 5**.

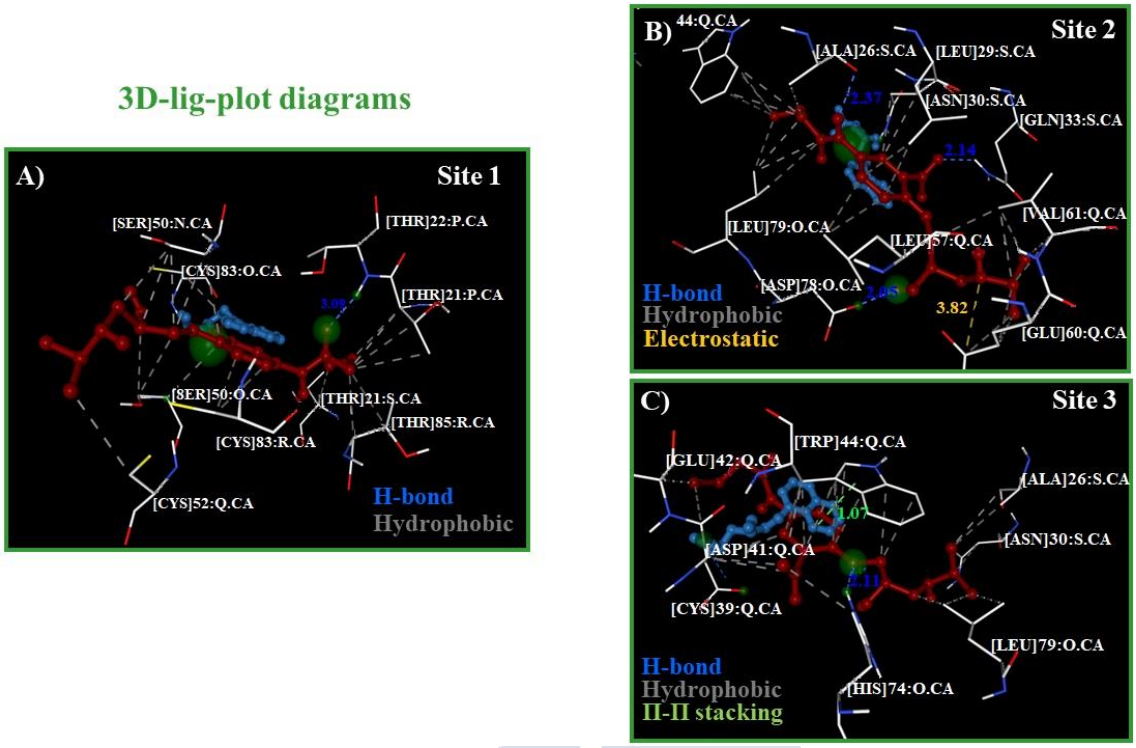


Figure 5. Representation of 3D-lig-plots diagrams for the best beta-blocker drug–drug binding-poses. **(A)** A plus P at site 1; **(B)** A plus P at site 2; and **(C)** A plus P at site 3 (A displayed in red color and P in blue). Therein, the atoms from target-residues colored in green correspond to potential hotspots for maximum drug–drug interactions between both beta-blockers in each fibrinogen binding site evaluated.

As can be seen, the A molecule involves a greater number of target-residues compared to P, covering a greater area in the biophysical environment shared by both beta-blockers during the docking interactions in the three binding sites. In all the cases, the best drug–drug binding pose (A plus P) showed an extended shape to interact with the three fibrinogen sites. Additionally, considering the approximate lengths of both molecules (A: ≈16.1 Å , and P: ≈ 4.8 Å), this suggests that the pharmacodynamic DDIs are mostly influenced by the contribution of the atoms from beta-blocker A, since it displayed a greater number of hydrophobic contacts with binding residues following the "mutual overlap criterion" cited above. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

An overview of the results revealed that the most relevant beta-blocker DDIs are the hydrophobic (N···C···C)-backbone side chain non-covalent interactions [46,47]. A greater number of these pertain to binding site 1 when compared to sites 2 and 3, which show a similar proportion. Thus, the presence of multiples hydrophobic (N···C···C)-

backbone side chain non-covalent interactions between the beta-blockers (A > P) with the fibrinogen binding sites promote thermodynamic destabilizing effects on the residue site-chain packing [46,47], inducing local perturbations in the flexibility and ligand-binding properties of those binding sites (site 1 > site 2 > site 3). Furthermore, other detected relevant interactions for these beta-blockers are an electrostatic (attractive interaction) between the electron-acceptor (N+)-(A) atom with two (O−)···δ-GLU60: chain Q (site 2) and a strong hydrophobic π···π stacking interaction between the naphthalene moiety belonging to P with the aromatic heterocyclic N-indole moiety-TRP44: chain Q (site 3), which can influence the stabilization of the A molecule that shares the same binding residue (TRP44: chain Q) through hydrophobic interactions with the benzene ring that is part of that aromatic heterocyclic N-indole moiety. Another aspect deserving special attention for targeting the beta-blocker DDI is the formation of a stable multicomponent docking complex that follows a recognized heterogeneous Monod-Wyman-Changeux Allosteric Transition model (HMWC-AT) [48–50], which depends on the fractional occupancy parameter (δ_{site}) as a measure of the affinity constant (K_i) or degree of saturation from the different binding sites [48–50]. The HMWC-AT model is related to the cases where two or more ligands interact with different binding sites of a given target-protein, but where just one affinity (ΔG_{bind}) per ligand is attributable [49]. In this context, the fractional occupancy for a given fibrinogen binding site forming a multicomponent complex with two different ligands (A and P) can be defined per binding site as follows in the **equation 1**:

$$\delta_{\text{site}} = \frac{1}{n}\left\{ \frac{\left[\sum_{i=1}^{n_1}(|\,\mathrm{A}\,|\,K_i^R\prod_{j\neq i}(1+|\,\mathrm{A}\,|\,K_j^R))\right]\times\left[\sum_{i=1}^{n_2}(|\,\mathrm{P}\,|\,K_i^R\prod_{j\neq i}(1+|\,\mathrm{P}\,|\,K_j^R))\right]}{\prod_{i=1}^{n_1}(1+|\,X_1\,|\,K_i^R)\prod_{i=1}^{n_2}(1+|\,X_2\,|\,K_i^R)+L\prod_{i=1}^{n_2}(1+c_i\,|\,X_1\,|\,K_i^R)\prod_{i}^{n_2}(1+c_i\,|\,X_2\,|\,K_i^R)} \right.$$

$$\left. + \frac{L\left[\sum_{i=1}^{n_1}(c_i\,|\,\mathrm{A}\,|\,K_i^R\prod_{j\neq i}(1+c_j\,|\,\mathrm{A}\,|\,K_j^R))\right]\times\left[\sum_{i=1}^{n_2}(c_i\,|\,\mathrm{P}\,|\,K_i^R\prod_{j\neq i}(1+c_j\,|\,\mathrm{P}\,|\,K_j^R))\right]}{\prod_{i=1}^{n_1}(1+|\,\mathrm{A}\,|\,K_i^R)\prod_{i=1}^{n_2}(1+|\,\mathrm{P}\,|\,K_i^R)+L\prod_{i=1}^{n_2}(1+c_i\,|\,\mathrm{A}\,|\,K_i^R)\prod_{i=1}^{n_2}(1+c_i\,|\,\mathrm{P}\,|\,K_i^R)} \right\} \quad \textbf{(1)}$$

In this equation, the first ligand (Acebutolol: A) interacts with the (n_i)-fibrinogen binding sites (n_1: sites 1, 2, or 3), and the second ligand (Propranolol: P) simultaneously interacts with the same n_i fibrinogen binding sites (n_2: sites 1, 2, or 3), but with different affinity (K_i or K_j) and both with the best D_g. Furthermore, X_1 and X_2 denote the ligand concentration of the evaluated beta-blockers, and the parameter L is an isomerization

constant of the fibrinogen E-region formed by Aα and Bβ chains with pseudo-allosteric behavior. This parameter is obtained by the ratio: $L = |T_0|/|L_0|$, which describes the equilibrium between the high-affinity relaxed conformation state (R_0) and the low-affinity tensor conformation state (T_0) for the three best-ranked fibrinogen binding sites [48–51]. In this context, the different affinity (K_i) is represented in the high-affinity R-state like K_i^R for both beta-blockers, and the c parameter indicates how much the equilibrium between T and R states changes under beta-blocker interactions [50–52]. The value of $D_g \approx \Delta G_{max}$ (Fibrogen site) is a function directly associated to the fractional occupancy parameter (δ_{site}), namely: $\delta_{site\ 1} \approx 0.81 > \delta_{site\ 2} \approx 0.54 > \delta_{site\ 3} \approx 0.39$. Let us now begin by ascertaining the total binding affinity (ΔG^T_{bind}), as well as its thermodynamic contributions (ΔG_{Gauss1}, ΔG_{Gauss2}, $\Delta G_{repulsion}$, $\Delta G_{H\text{-}bond}$, $\Delta G_{hydrophobic}$, and ΔG_{rot}; see **equations (11)** and **(12)** in the **Materials and Methods** section). Additionally, the frequencies (k-occurrences) of affinities were evaluated by assuming that the ligand binding occurs in the high-affinity relaxed conformation state (R_0) in the three best-ranked binding sites (site 1 > site 2 > site 3). The results of this analysis are shown in **Figure 6**. An analysis of the results depicted in **Figure 6** revealed that the most dominant energy contributions for the total binding affinity ΔG^T_{bind} of the beta-blocker DDI systems are the pair of steric Gaussian molecular mechanism distance-dependent terms (Gauss2 > Gauss1) [53]. In particular, the Gauss2 energetic terms provide relevant information about the influence of attractive binding energies of the beta-blockers system with the three fibrinogen binding sites (site 1 > site 2 ≈ site 3). These show the presence of the strongest interactions based on non-covalent hydrophobic and non-electrostatic attractive energy contributions between the A and P hydrophobic atoms making the DDI system. This indicates that the great influence exerted by the Gauss2 energy [53] is due to the formation of temporary dipoles (fluctuating dipole–induced dipole) with electron cloud overlap between both beta-blockers. It is well-known that this phenomenon can induce random fluctuations of electron density and non-zero instantaneous dipole moments in all the interacting atoms, therefore leading to a significantly increase in the total affinity when both beta-blockers interact at the same biophysical environment in the fibrinogen E-region binding site. Particularly, based on the Gauss2 attractive energy contributions, the narrowest pocket seen in the case of the binding site 1 appeared to favor a steeper stabilization of the drug-drug system when compared to binding sites 2 and 3. Indeed, the later sites presented a greater number of tunnels and binding interaction surfaces that decrease the probability of drug–drug interactions based on their frequency (k-

occurrences). These conclusions match up with the previous results obtained by the ligand transport approach [54], where very close values of maximum affinity were found for sites 2 and 3 compared to the first-ranked binding site 1. In addition, the hydrogen-bond interactions contribute to the stabilization of the beta-blocker drug-drug complexes in the system but show less impact on its total affinity, which display similar interaction values for the three catalytic active sites studied. On the other hand, from a quantitative point of view, the energetic contributions related to repulsive and rotational forces showed a similar behavior for the three binding sites. So, repulsive and rotational forces are thought to have low relevance for the stabilization of the beta-blocker drug-drug complexes in the fibrinogen E-region binding sites evaluated. Furthermore, the energetic contributions of the beta-blockers A and P with the "worst fibrinogen binding site" were used as a control simulation experiment for comparison purposes (see **Figure S3**).

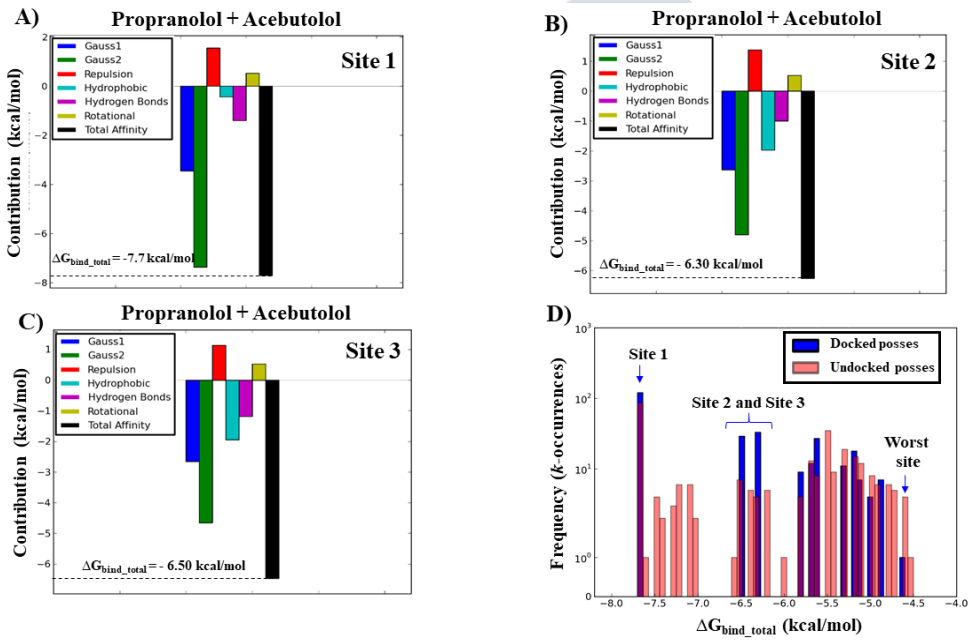


Figure 6. Graphical breakdown of the different binding energy contributions of the total binding affinity ($\Delta G^{T}_{\text{bind}}$) with the corresponding values obtained from the simultaneous interaction of both beta-blockers with the fibrinogen E-region binding sites. **A)** A and P at site 1 ($\Delta G^{T}_{\text{bind}}$ = □7.70 kcal/mol); **B)** A and P at site 2 ($\Delta G^{T}_{\text{bind}}$ = □6.30 kcal/mol); and **C)** A and P at site 3 ($\Delta G^{T}_{\text{bind}}$ = □6.50 kcal/mol). Therein, different colors were used to distinguish the type of thermodynamic contribution, namely: attractive energy dispersion based on two Gaussian functions (Gauss1 and Gauss2 in blue and green, respectively), repulsion energy (red), hydrophobic energy (light blue), hydrogen bond energy (purple), rotational energy (yellow) and total affinity (black). **D)** Frequency distribution of the total binding energies, in which the *x-axis* pertains to the drug-drug total affinities ($\Delta G^{T}_{\text{bind}}$) and the *y-axis* to their frequencies of occurrence (*k*-occurrences),

obtained from both the docked drug-drug poses (blue bars) and the undocked drug-drug poses (pink bars). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

On the other hand, the results shown in **Figure 6D** corroborate that the beta-blocker DDI events present a higher prevalence in the fibrinogen E-region site 1 compared to sites 2 and 3, based on counting the binding affinity frequencies (k-occurrences) [54]. In fact, the calculated values of k-occurrences in site 1 (k = 100) were significantly higher compared to those in sites 2 and 3 (k = 10–13), suggesting the high propensity of the fibrinogen E-region binding site 1 for DDI events [53] To complement the previous analysis, we carried out additional control simulation experiments in order to identify which specific atom, or set of atoms, of the beta-blockers could present the maximum energy contribution to the individual binding affinities (ΔG A bind or ΔG P bind), taking into account again the same thermodynamic energy contributions. To this end, a per atom energy contribution analysis was performed separately for each beta-blocker (see **Figure 7**). In fact, such analysis allows a more detailed evaluation of the contribution of the individual atoms of the beta-blockers interacting separately with separately with the fibrinogen E-region binding sites [33,34,41,42,55,56].

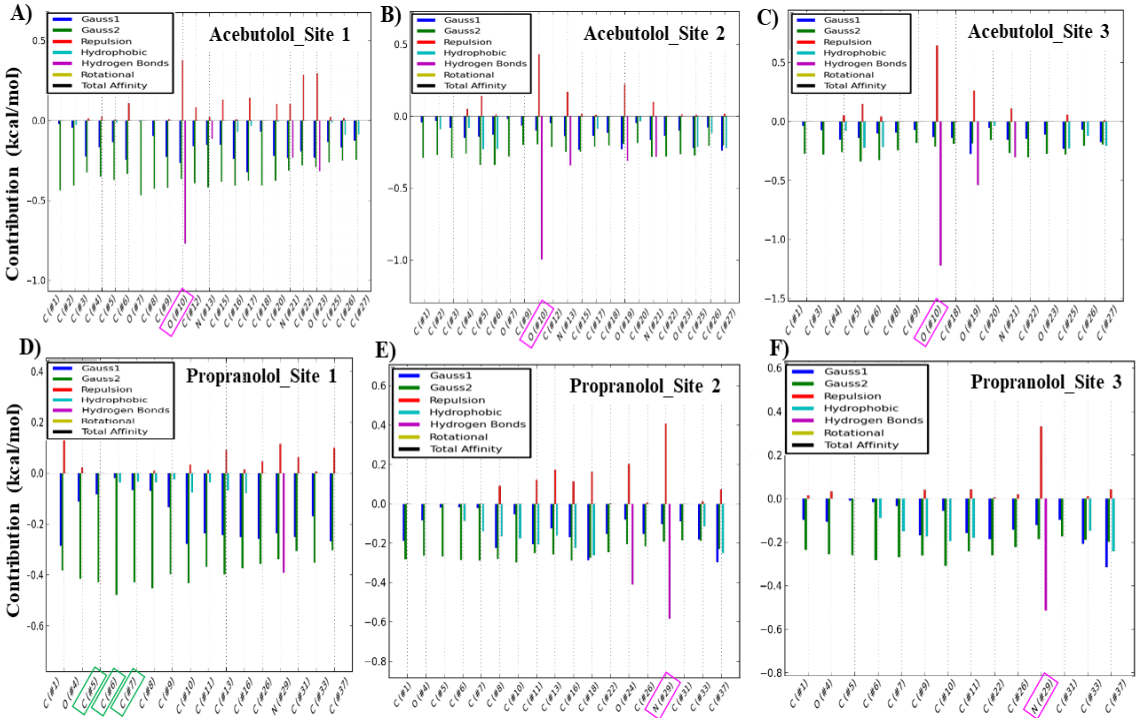


Figure 7. Graphical representation based on the beta-blocker *per atom* energy contributions to the individual binding affinity (ΔG^{A}_{bind} or ΔG^{P}_{bind}). in the three best-ranked binding sites. On the top, energy contributions for each atom of A interacting with the

three best-ranked binding sites: **A)** A-oxygen-(2)-atom/site 1, **B)** A-oxygen-(2)-atom/site 2, and **C)** A-oxygen-(1)-atom/site3. On the bottom, energy contributions for each atom of P interacting with the three best-ranked binding sites: **D)** P-C-atoms (C2, C3, C4)/site 1, **E)** P-N-atom/site 2, and **F)** P-N-atom/site 3. Please, note that therein the symbol (#) *x-axis* is only used for labeling purposes and do not fit with the atomic position in the beta-blocker molecular structures. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

As can be observed, for molecule A, the largest energetic contribution was given by the H-bond interaction generated by the A (H)···O-atom-(#2), with an oxo-group with electron-acceptor properties that interacted with the (H)-THR22: chain P:CA in binding site 1 (**Figure 7A**). In site 2, the highest energetic contribution for molecule A was supplied by means of H-bond interaction between the (H)-axial-(#2) of the (OH)-hydroxyl group of A, which exhibited electron-donor properties to interact with the (H)···ASP78: chain O: CA (**Figure 7B**). In the case of site 3, the atomic energy contribution was provided by the H-bond interaction between the (H)···O-atom-(#1)-A, which interacted with the (H)-axial-(#2) of the (OH)-HIS74 hydroxyl group in a similar way to the previous interaction in chain O (**Figure 7C**). The maximum per atom energetic contributions for the A molecule to ΔG_A bind were then provided by hydrogen bond interactions, these being remarkably close in the three fibrinogen E-region binding sites (ΔGH-bond ≈ −0.8 kcal/mol).

Regarding the P molecule, the maximum per atom energy contribution to the individual ΔG_P bind was provided by a triad of C-atoms: C-atom (#2) > C-atom-(#1) ≈ C-atom-(#3) in order of influence, with the SER50: chain Q:CA interacting through the hydrophobic (N···C···C)-backbone and site chain (based on the attractive Gauss2 energy of ca. −0.5 kcal/mol) in site 1 (**Figure 7D**), and also inducing hydrophobic clashes. In binding site 2, the most relevant atomic energy contribution of the P molecule was provided by means of ΔGH-bond interaction between the electron-acceptor (H)···N-acylated-atom-(#1) of P, which interacted with the (H)···N-atom of the ASN30: chain S:CA (**Figure 7E**). Lastly, in site 3, the most significant energetic contribution to ΔG P bind was led by ΔGH-bond involving the same electron-acceptor (H)···N-acylated-atom-(#1) in the P molecule, which interacted with the (H)-atom linked to the (H)···O-CYS39: chain Q: CA (**Figure 7F**). These two lasts per atom energetic contributions of the P molecule interacting with sites 2 and 3 exhibited remarkably close values of ΔGH-bond (≈ −0.5 kcal/mol).

Additional details such as control simulation experiments for the "worst fibrinogen binding site" are provided in **Figure S4** of the Supplementary Material.

Let us now assess the perturbations in the depth profiles of target residues making relevant tunnels (THR22:P, SER50: Q/site1, ASP78:O, ASN 30:S/site2 and HIS 74:O, CYS39: Q/site3) that interacted with the beta-blocker ligand atoms previously identified, with the maximum per atom energy contribution to the individual binding affinities [39,40,44] (see **Figure 8**).

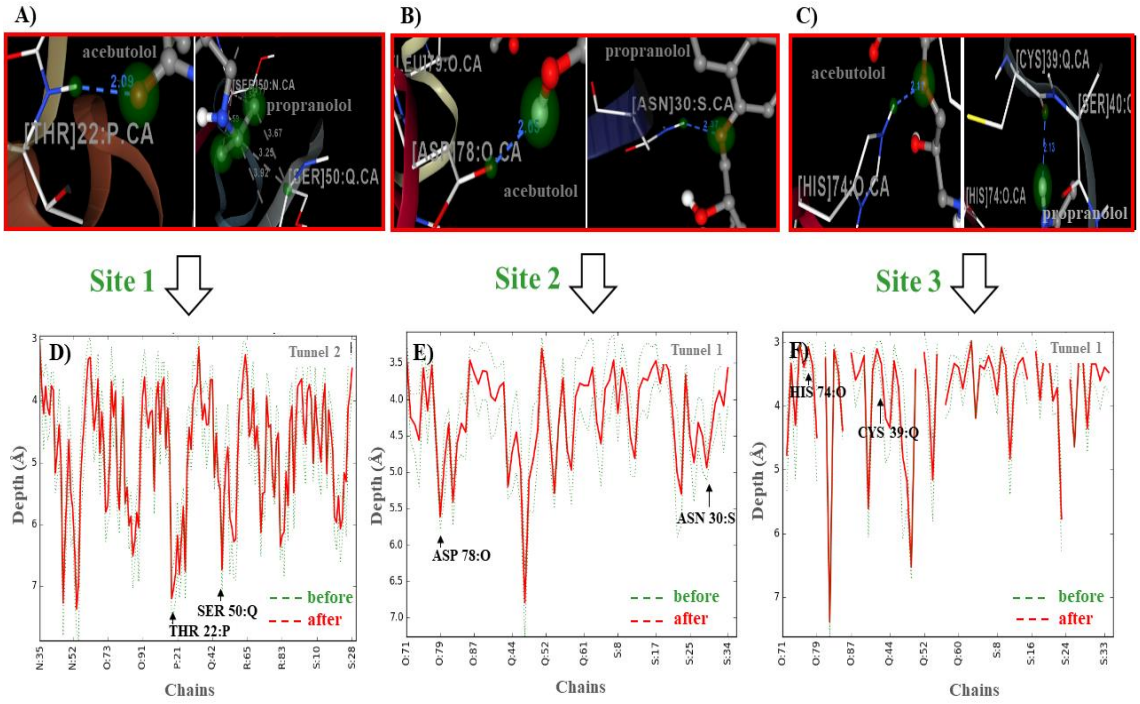


Figure 8. On the top, for each beta-blocker, critical target-residues interacting with the ligand atoms that showed the maximum atom energy contribution to the individual binding affinity. (A) THR22:P_site1/(A)-oxygen-(2)-atom and SER50_site1: Q/(P)-C-atoms_(C2, C3, C4); (B) ASP78:O_site2/(A)-oxygen-(2)-atom and ASN30:S/(P)-N-atom; (C) HIS74:O_site3/(A)-oxygen-(1)-atom and CYS39: Q_site3/(P)-N-atom. On the bottom, (D–F) correspond to the depth profiles (D[c,i] vs. chains) for the critical target-residues linked to their corresponding side-chains into the relevant tunnels. Therein, the amplitude of the depth peaks was evaluated before (green dashed lines) and after beta-blocker interactions (red solid lines). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

According to the results shown in **Figure 8** (panels: **D, E,** and **F**), the relationship between the pharmacodynamics drug-drug interactions and the perturbations on the depth values for the revealed target-residues of A-A and P-P could be modeled by incorporating the latter into previous eq 6, leading to expressions for $D_{[\mathrm{THR22}\cdot\cdot\mathrm{SER50}]}$, $D_{[\mathrm{ASP78}\cdot\cdot\mathrm{ASN30}]}$, and $D_{[\mathrm{HIS74}\cdot\cdot\mathrm{CYS39}]}$. In those expressions, it appears the q_i^c descriptor (*i.e.*: q_{THR22}^P, q_{SER50}^Q, q_{ASP78}^O, q_{ASN30}^S, q_{HIS74}^O, and q_{CYS39}^Q), derived from the distance matrix between the 20 standard amino acids, which depends on the frequency of the target residues in a certain position for a given chain (*c*). As referred to before, this descriptor might be used as an adjusted residue-binding probability to establish the *quasi-sequence-position-order descriptor* $^{J_i}I_i$ [39, 40, 44], which depends of the position (*i* and *i* + *d*) and the frequency-based occurrence of the target residues in the chain sequence. In this context, for each target residue type, the depth descriptor q_i^c is defined by means of the following **equation**:

$$q_{i\,(J_i)}^{c} = \frac{f_r}{\sum_{i=1}^{20} f_r + w \sum_{d=1}^{L} \tau_d} \quad | \, i_{\text{-target}} = 1, 2, \ldots, 20 \tag{2}$$

in which τ_d is defined as follows:

$$\tau_d = \sum_{i=1}^{N-d} \left(d_{i,[i+d]}\right)^2 \quad | \, d = 1, 2, \ldots, L_{\max} \tag{3}$$

In Equation (2), f_r stands for the normalized frequency-based occurrence of the target residue analyzed (i.e., THR22: P, SER50: Q/site 1, ASP78:O, ASN30:S/site 2 and HIS74:O, CYS39: Q/site 3), and w is a weighting factor (w = 0.1). In Equation (3), the parameter τ_d stands for the d-th rank sequence-order-coupling descriptor based on the C-(α)-distances ($d_{i,[i+d]}$) between the two neighbors most contiguous residues around the target-residue at the position (i, $i + d$), and $L_{\max}$ for the maximum length of the chain (c),

where the fibrinogen target-residues are included. Then, one can estimate the residue depth-perturbations after the beta-blocker interactions interactions (J_i^c) by correcting this descriptor with an adjusted residue-binding probability under the influence or binding perturbation with the beta-blocker ligands (lig $_{\beta\text{-blocker}}$: A and P) *i.e.,* $J_i^c + lig_{\beta-blocker}$: $J_i^c + lig_{\beta-blocker}$: $J_i^c + lig_{\beta-blocker}$: site 1: $J^P_{\text{THR22/A}}$, $J^Q_{\text{SER50/P}}$ $J^P_{\text{THR22/A}}$, $J^Q_{\text{SER50/P}}$; site 2: $J^O_{\text{ASP78/A}}$, $J^S_{\text{ASN30/P}}$; and site 3: $J^O_{\text{HIS74/A}}$, $J^Q_{\text{CYS39/P}}$. From the results shown in **Figure 8** (panels **D**, **E**, and **F**), it can be seen that, in general terms, the beta-blocker drug–drug interactions induced a decrease in the residue depth values (D[c,i]) of approximately 1-2 Å for the target-residues analyzed. Such results, therefore, suggest that the simultaneous interactions of both beta-blockers can induce loss of coupling between the target-residues (i.e., changes in the parameter τ_d), leading to anisotropic changes in the C-(α)-distances of the interacting target residues from their equilibrium coordinates, and depth-perturbations (i.e., displacements) measured by the ($D_{[c,\,i]}$) parameter. Additionally, these molecular changes could significantly affect the substrate-specificity for thrombin the natural ligand substrate of the fibrinogen E-region, and flexibility properties of the fibrinogen binding sites [39,40,44] (see **Figure 9**).

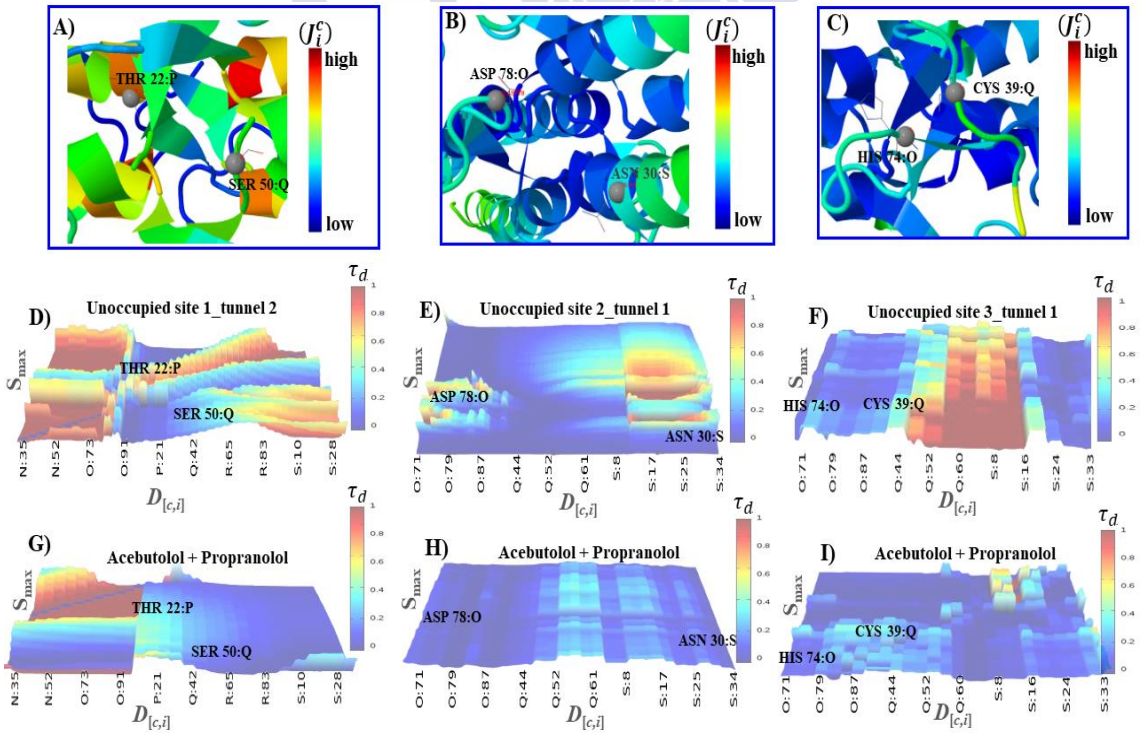


Figure 9. On the top, graphical representation of the beta-blocker binding ability for critical fibrinogen target-residues based on the depth quasi-sequence-position-order descriptor (J_i). (A) THR22:P, SER50: Q; (B) ASP78:O, ASN30:S; and (C) HIS74:O, CYS39:Q. Orange and red colored regions depict high binding ability for the beta-blockers, while blue ones show low binding ability. On the bottom, 3D-surface maps of the quasi-sequence-order-coupling descriptor (τ_d) as a function of residue depth D[c,i] vs. maximum solvent accessibility. The simulation conditions considered to study the depth perturbation on the aforementioned target-residues are: (D) unoccupied site 1 (tunnel 2); (E) unoccupied site 2 (tunnel 1); and (F) unoccupied site 3 (tunnel 1). (G-I) The later panels refer to the last simulation condition considered that is, the study of the residue depth perturbations of the aforementioned target-residues under the influence of drug–drug interactions with both beta-blockers A plus P. Therein, the color bar on the right of each map represents the residue depth perturbations based on the values of descriptor τd, which range from weak (blue) to strong (orange/red) ones. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

As can be seen, the fibrinogen site 1 led preferably to beta-blocker DDIs considering that the modifications in the residue depth profiles (THR22:P, SER50: Q) were more pronounced compared to those on site 2 (ASP78:O, ASN30:S) and site 3 (HIS 74:O, CYS 39: Q) after these occurred. Furthermore, considering the different fibrinogen target-chains c, one can see that tunnel 2 at fibrinogen site 1 was the most preferred biophysical environment for the beta-blocker drug–drug interactions. Indeed, tunnel 2 at site 1 showed the highest relevance involving all the fibrinogen target-chains (N, O, P, Q, R, S) with c[site 1] = 6 compared with that at tunnel 1 of sites 2 and 3 that display medium relevance (c[site 2] = c[site 3] = 3) and involving only the chains O, Q, and S. Additionally, the occurrence of the beta-blocker DDI phenomena was more likely at site 1 than sites 2 and 3 since it obeys a tighter relationship with D_g, S_{max}, and $D_{[c,i]}$ (see **Figure S5**). From the polypharmacology point of view, these results have a remarkable value since they match with previous in vitro evidence, which suggest that modifications in the residue depth profile can significantly affect several binding properties of the catalytic sites such as substrate-specificity, enantioselectivity, inhibition constants (K_i), and thermodynamic-stability of the protein–ligand complexes [39,40].

5.2.4. DFT Modeling of Beta-Blocker Drug–Drug Binding Systems

In order to further characterize the structural and electronic properties of the most stable beta-blocker drug-drug binding configurations, a DFT study was then carried out [57]. We started by evaluating the structural and electronic properties of the isolated beta-blockers A and P (see **Figure 10**).

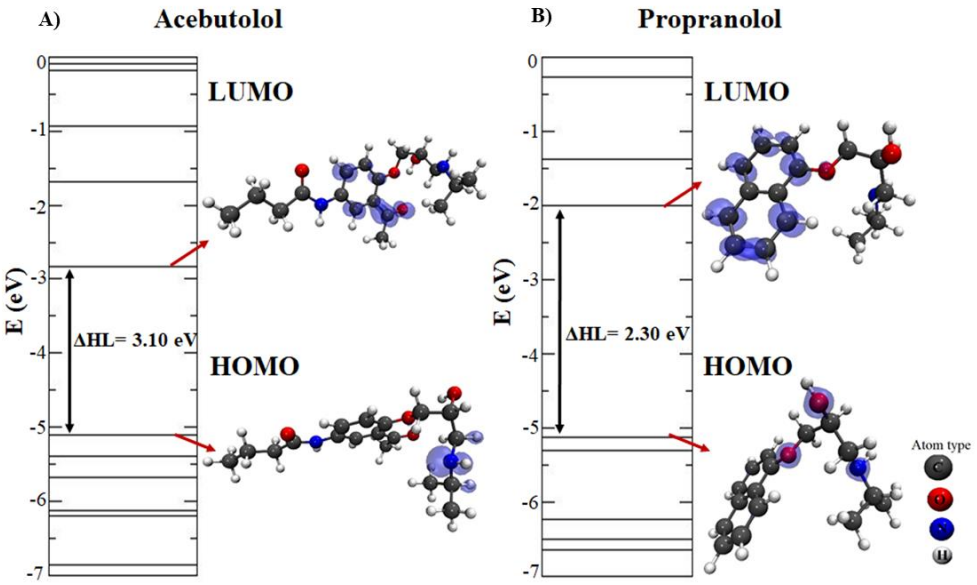


Figure 10. Representation of the electronic energy levels and local charge density plots with the corresponding HOMO and LUMO levels (4.3×10^{-3} e/Å^3 iso-surface) for the optimized beta-blocker ligands obtained by DFT calculations. **A)** beta-blocker A and **B)** beta-blocker P. In the lower right corner, the atom types of the beta-blockers are depicted, namely: C-atoms (dark gray), O-atoms (red), N-atoms (blue), and H-atoms (light gray). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

To do so, we have calculated the electronic levels of the isolated beta-blockers with the corresponding local charge densities plots for the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) [58, 59]. Here it should be noticed that local charge density plots allow us to identify relevant target-atoms potentially involved in the beta-blocker DDI from an electronic point of view [58, 60]. For molecule A, the obtained HOMO–LUMO energy gap difference (ΔHL) was 3.10 eV, and, as it can be seen, the density of charges in the HOMO level is essentially placed in the N-atom, while in the LUMO level those are more concentrated in the (sp^2)-C-atoms of the aromatic ring, including the O-atom from one part of this molecule. On the other hand, the obtained HOMO-LUMO gap in the case of molecule P was 2.30 eV, with the

HOMO level more concentrated in the N-atom and O-atom, while its LUMO level covering the (sp^2)-C-atoms from the naphthalene-moiety.

Then, we carried out DFT calculations for all the potential conformations adopted by the beta-blockers during the pharmacodynamics interactions, according to the ligand binding conformations and thermodynamic stability. To do so, the most stable drug-drug binding configurations of the beta-blockers were identified following several criteria for their prioritization based on the following DFT obtained results: drug-drug energy of binding (E_b), drug-drug inter-atomic distances of binding ($d_{A\text{-}P}$) for the molecules A and P, and the HOMO-LUMO energy difference gap (ΔHL) of each modeled system (see **Table 1**). In so doing, fifteen most stable drug-drug binding configurations from a total of forty modeled configurations were found (see **Figure 11**).

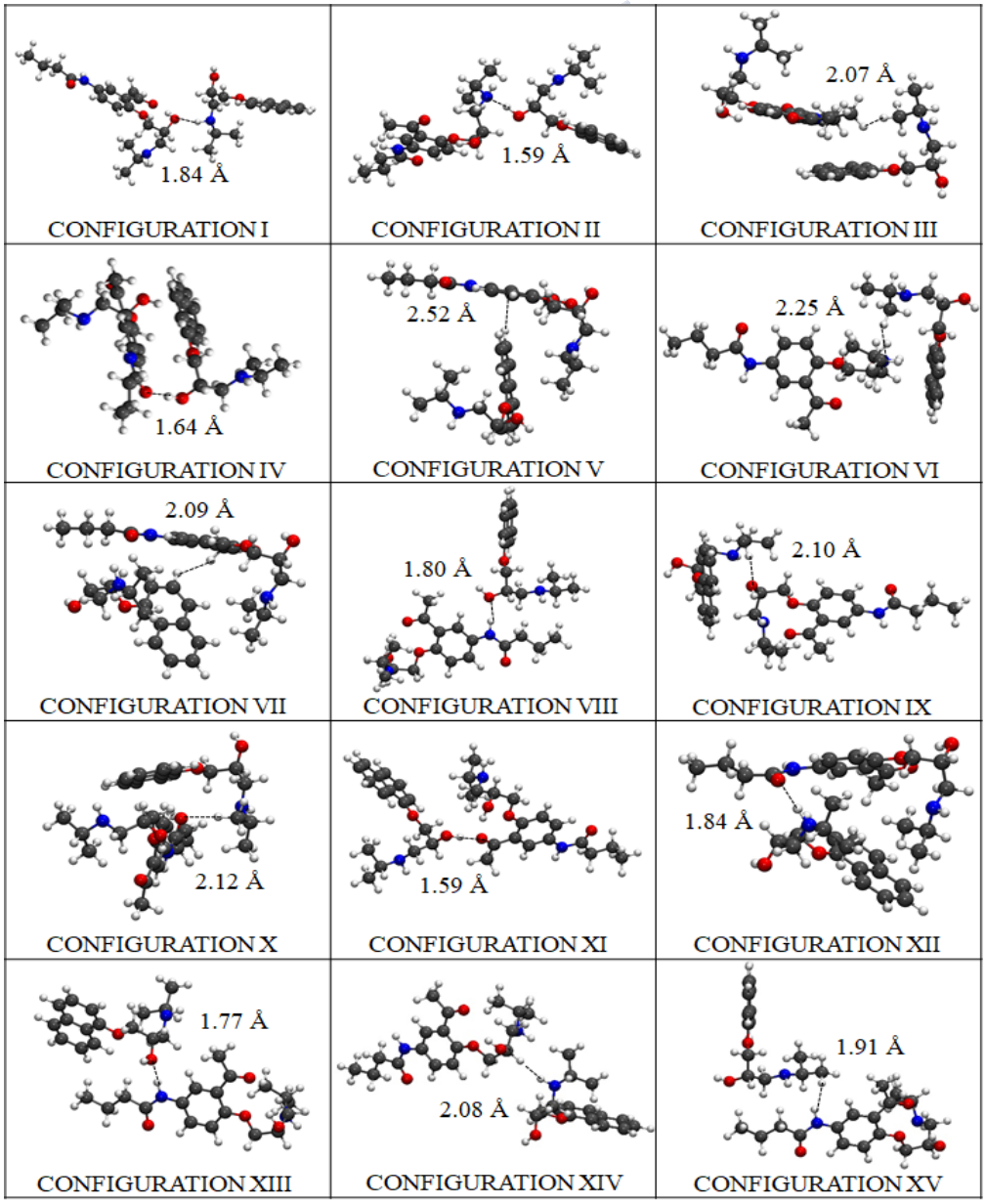


Figure 11. DFT results for the fifteen most stable thermodynamically drug-drug binding configurations (from I to XV) adopted by the beta-blockers A and P during the pharmacodynamic interactions (*i.e.*, modeled in the absence of fibrinogen E-region). The beta-blocker atoms are depicted as ball and sticks and colored as follows: dark gray (C-atoms), red (O-atoms), blue (N-atoms), and light gray (H-atoms). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

Table 1. Energy of binding (E_b), lowest β-blocker inter-atomic distance of interaction (d_{A-P}), and HOMO-LUMO gap (ΔHL) obtained for the fifteen most stable configurations of beta-blocker DDI systems ordered according to their pharmacodynamics relevance (DFT results obtained by modeling DDI in the absence of fibrinogen E-region molecule). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

Ranking	Configuration	$\|E_b\|$ (eV)	ΔHL (eV)	d_{A-P} (Å)
1	XII	2.40	2.02	1.84 (O-H)
2	IV	2.36	2.11	1.64 (O-H)
3	II	1.99	1.87	1.59 (N-H)
4	XIII	1.99	2.26	1.77 (H-O)
5	XIV	1.99	2.15	2.08 (N-H)
6	XI	1.97	1.63	1.59 (O-H)
7	VI	1.92	1.90	2.25 (N-H)
8	VII	1.88	2.13	2.09 (H-H)
9	X	1.84	2.29	2.12 (O-H)
10	VIII	1.83	2.30	1.80 (H-O)
11	III	1.82	2.22	2.07 (H-H)
12	I	1.76	1.47	1.84 (O-H)
13	IX	1.71	2.20	2.10 (O-H)
14	V	1.67	1.97	2.52 (C-H)
15	XV	1.67	2.21	1.91 (H-H)

The binding energy $|E_b|$-based distance is considered the most relevant DFT result to tackle the DDI between both beta-blockers. It is important to note here that, the DDI phenomena strongly depend on the best thermodynamic stability of the configuration

binding-poses adopted by the beta-blockers in the binding site 1. Regarding this, the DFT results point out that the best-ranked configuration binding-pose is achieved by configuration XII, which presents the highest drug-drug binding-energy value $|E_b|$ = 2.40 eV, with an HOMO-LUMO difference gap of ΔHL = 2.02 eV associated to an inter-atomic distance of interaction ($\boldsymbol{d}$(O···H)$_{A\text{-}P}$) = 1.84 Å, established between the O-atom of molecule A and the H-atom of molecule P. Short inter-atomic distances such as O···H between molecules A and P in this configuration favor the non-covalent stabilization of the DDI system, forming a (O···H)-dipole induced-dipole interaction without involving the sharing of electrons. Details on this analysis are shown in **Figure 12.**

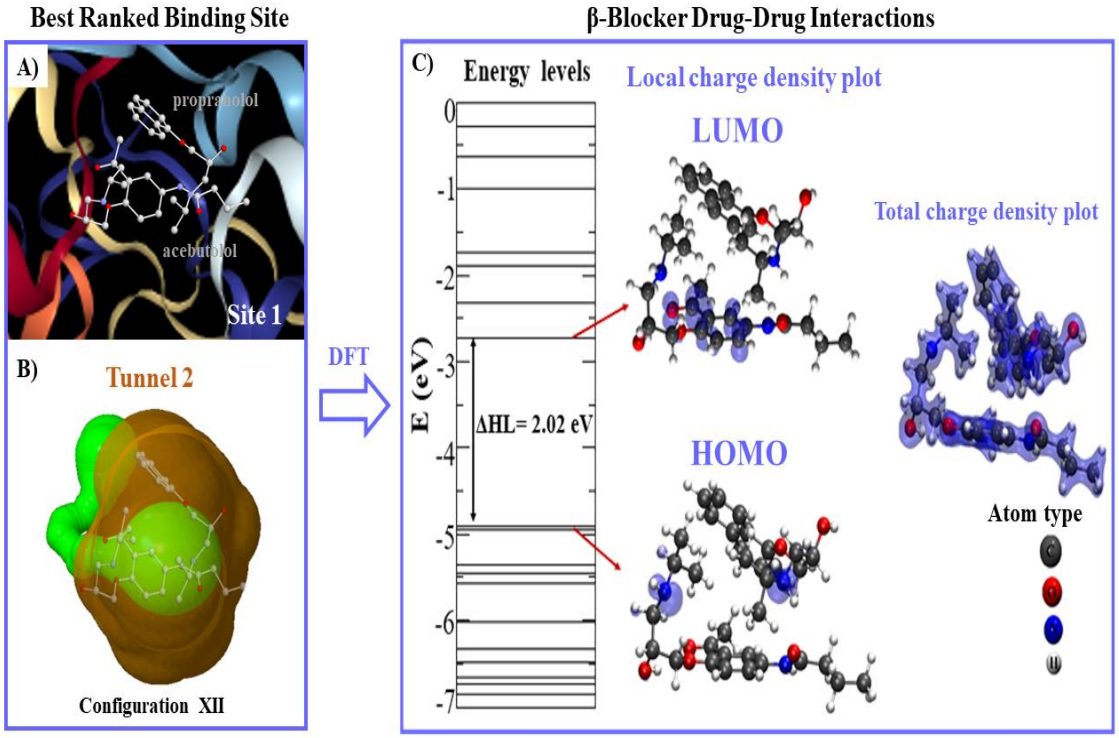


Figure 12. **A)** Representation of the best-ranked drug-drug binding configuration XII formed by the A + P/fibrinogen complex in the fibrinogen E-region binding site 1. **B)** Topological representation of the tunnel 2 (colored green) containing the best-ranked drug-drug binding configuration XII for molecules A + P interacting simultaneously at the same biophysical environment (ball and sticks representation) in binding site 1 (surrounding shaded orange region). **C)** Energy levels, and iso-surfaces of the local charge electronic density and total charge density plots obtained for the best-ranked drug-drug binding configuration XII (A + P). The electronic levels are depicted as solid lines with the HOMO-LUMO energy difference gap (ΔHL), and the density plots are shown, using orbital charge density iso-surfaces of 0.002 eV/Å^3. The A and

P atoms are depicted as ball and sticks and colored as follows: dark gray (C-atoms), red (O-atoms), blue (N-atoms), and light gray (H-atoms). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

These DFT results go along with the previous docking results, since configuration XII is part of the clustered docking poses that exhibit higher frequency values for DDI in the fibrinogen E-region site 1 (*i.e.*, at site 1 *k* occurrences = 100) as referred to before (see **Figure 6D**). Also, configuration XII exhibits a greater structural overlap between the aromatic moieties (*i.e.*, benzene and naphthalene) compared with the remaining configurations (**Figure 11**). Interestingly, the second-best ranked drug-drug binding configuration corresponds to configuration IV, which present a slightly lower-energy value ($|E_b|$ = 2.36 eV) compared to the former one (XII), but in contrast, it shows a lower value for the inter-atomic distance O···H ($\boldsymbol{d}(\mathrm{O}\cdots\mathrm{H})_{\boldsymbol{A\text{-}P}}$ = 1.64 Å). Besides, the slightly difference of 0.04 eV obtained for configuration IV in relation to the best drug-drug binding configuration XII ($|E_b|$ = 2.36 eV) may be explained by the presence of structural overlapping parallel geometry between the aromatic benzene ring of A and the naphthalene-moiety of P. This clearly indicates potential influence of the π-π stacking in the stabilization of the beta-blocker DDI in this case.

In this context, by analyzing the worst-ranked configuration XV with $|E_b|$ = 1.67 eV, one can observe a clear separation of the just referred aromatic-moieties, which are oriented in a different plane (please refer to **Figure 11**). Also, the presence of β-blocker inter-atomic distance of interaction $d_{A\text{-}P}$ (H-H) = 1.91 Å and HOMO-LUMO difference ΔHL = 2.21 eV could negatively influence the stabilization of configuration XV, and thus decrease the DDI phenomena in the fibrinogen E-region site 1 compared to the best-ranked drug-drug binding configurations (XII > IV) (please refer to **Table 1**).

Considering the electronic properties for the best-ranked configuration (XII) in the fibrinogen E-region (binding site 1), one can note locally concentrated charges in the LUMO-orbital of A mainly placed on the (sp^2)-C-atoms of the aromatic benzene-ring and covering partially the O-atom that forms the *oxo*-moiety in the A molecule. This fact, could significantly weaken the H-bond electron-acceptor properties of the *oxo*-moiety that interacts with the (H)-atom of the regulatory residue THR22: chain P:CA in the fibrinogen binding site 1. Therefore, in this way, it will affect the stabilization of the A-fibrinogen docking complex as well as significantly the pharmacological activity of the A, when beta-blocker P interacts in the same biophysical environment (site 1). Regarding

the HOMO-orbital, the local charge density of configuration XII is spread on the N-atoms simultaneously for both beta-blockers. However, the latter does not involve relevant interacting-atoms for stabilizing the (A + P)/fibrinogen complex in site 1. On the other hand, the results concerning the total charge density plots [52] show that the charges are localized just on the separated beta-blockers without charge lines intercepting both molecules, thus confirming the absence of DDI under a covalent regime between the beta-blockers A and P. Also, the obtained DFT results obtained for the binding energies ($|E_b|$) for all the configurations of (A + P) were smaller than 2.50 eV indicating that the interactions occur without chemical bonds [58, 59]. Besides, the total charge density plots show that the geometric structure and inter-atomic bond-distances are similar to the ones of the original isolated molecules, suggesting that the beta-blocker DDI in site 1 could be more intense mimicking traditional pharmacodynamics DDI phenomena.

5.3 Materials and Methods

5.3.1 General workflow. To begin with, the protein receptor file - *i.e.*, fibrinogen E-region, was withdrawn from the Protein Data Bank [61], PDB ID: 1JY2 with 1.4 Å of resolution and more than 98% similarity with the human fibrinogen E-region [26]. Then, the E-region PDB structure was prepared using the AutoDock Vina software [62], by including titration of the protonation states using PROPKA 3.1 utility tool, followed by the addition of missing atoms with an overall optimization of the H-network employing its implemented PDB2PQR 2 algorithm. Next, the ligands (beta-blocker drugs) were obtained from the PubChem Substance and Compound databases [63], *i.e.*, A: hydrochloride molecule (PubChem CID:1978; MF: $C_{18}H_{28}N_2O_{24}$), and P: hydrochloride molecule (PubChem CID:62882; MF: $C_{16}H_{22}CL_2O_2$), and optimized using the semiempirical AM1 method implemented in MOPAC [64].

The following fifth-step procedure was adopted for addressing the study of the drug interactions between the beta-blockers A and P with the fibrinogen E-region binding-sites.

Step1: Prediction of the binding sites for fibrinogen. This proceeds as follows.

(i) Predict the binding sites for the fibrinogen E-region by applying an appropriate machine learning framework.

(ii) Select from the fibrinogen E-region binding sites found, the three best-ranked ones based on their highest druggability. For comparison purposes, ascertain as

well the undruggable or "worst fibrinogen E-region binding site" to be used as reference control denoting absence of drug-drug interactions.

(iii) Validate the three best-ranked binding sites found by examining their Ramachandran plots.

Step 2: Identification of tunnels for the fibrinogen binding sites. Detect the tunnels or tiny cavities of the three best-ranked binding sites found, followed by a ligand transport analysis for the beta-blockers A and P on those, and inspection of 3D interaction diagrams for their target-residues.

Step 3: Calculation of energetic contributions for binding affinity. Ascertain the energetic contributions for the total binding affinity from beta-blocker DDI systems on the three best-ranked fibrinogen E-region binding sites and evaluate in addition their binding relevance before as well as after the beta-blocker DDI.

Step 4: DFT modeling of beta-blocker drug-drug binding systems. Evaluate the configurations of the best-ranked beta-blocker drug-drug binding systems, using DFT calculations.

Step1: Prediction of the binding sites for fibrinogen. The task here is to predict the binding sites of the fibrinogen E-region, prior to the docking simulations. This was accomplished by applying the 3D DeepSite Convolutional Neural Networks (3D-DCNN) tool [32]. The latter considers all the crystallographic descriptors of a given binding site like van der Waals surface of the fibrinogen cavities using a validated deep-learning neural network algorithm. Herein, the characterization and ranking of the predicted binding sites were established by considering the ability of these sites to simultaneously bind both beta-blockers drugs A and P at the same biophysical environment [18-20].

Then, the cartesian coordinates obtained from the volumetric maps of the fibrinogen binding sites were used to setup the docking box simulation for each fibrinogen E-region binding sites. For site 1, the details of the box-simulation are: grid box-size with volume V= 559 $Å^3$ (X = 64 Å, Y = 52 Å, Z = 52 Å) centered at X = 13.6 Å, Y = -1.1 Å, Z = 12 Å; whereas for site 2, these are: grid box-size with volume V= 475 $Å^3$ (X = 24 Å, Y = 24 Å, Z = 24 Å) centered at X = 28.6 Å, Y = 2.8 Å, Z = 16 Å; and for site 3: grid box-size with volume V= 545 $Å^3$ (X = 28 Å, Y = 28 Å, Z = 28 Å) centered at X = 0.2 Å, Y= -8.4 Å, Z = 5.7 Å. Finally, for the "worst fibrinogen binding site", these are: grid box-size with volume V= 500 $Å^3$ (X = 28 Å, Y = 28 Å, Z = 28 Å) centered at X = 38.6 Å, Y = 13.1 Å, Z

= 31.1 Å. For this instance, an exhaustiveness docking parameter equal to 100 was used in all the cases [65].

To establish the relevance of each catalytic binding site, the druggability (D_g) was estimated as the maximum energy variation of the fibrinogen E-region binding-sites affinities under beta-blocker drug interactions. D_g is defined as follows in **equation 4**:

$$D_g \approx \Delta G_{max}(\text{Fib.site}) \approx -\gamma(r)\, A_{\text{Non-polarsite}} \times \frac{A_{\text{site-druglike}}}{S_{\text{max site}}} + C \quad \textbf{(4)}$$

where $\Delta G_{max}(\text{Fib.site})$ is similar to D_g [19], more commonly used in rational drug design to represent the maximum intrinsic affinity of the fibrinogen catalytic binding site, which fits to a linear-correlation between the binding-site surface area and the beta-blocker molecular weights under DDI conditions. Then, γ (=0.024 kcal/mol) stands for a constant related to solvent surface tension in the fibrinogen binding site, which depends on the surface curvature *r* (*i.e.*, concave surface) of the fibrinogen binding site. The $A_{\text{Non-polarsite}}$ corresponds to an approximation of the beta-blockers desolvation components, whereas the $A_{\text{site-druglike}}$ to the drug-like size (≈ 300 Å^2). The $S_{\text{max site}}$ depicts the normalized maximum solvent-accessible surface area of the fibrinogen site evaluated, and *C* is an empirical parameter that largely discriminates druggable from undruggable binding-sites when it is *ca.* zero [18-20].

Validation of the fibrinogen binding sites. To validate the 3D *X*-ray crystallographic structure of the predicted best-ranked fibrinogen E-region binding sites, namely: sites 1, 2, and 3, a Ramachandran diagram analysis (ψ *vs.* φ) was carried out [38]. Here it should be noticed that Ramachandran diagrams may prevent obtaining false positives from docking complexes, by confirming the absence of restricted-flexibility in the target-residues making up the fibrinogen E-region binding sites.

Step2: Identification of tunnels for the fibrinogen binding sites. Detection of tunnels for each fibrinogen E-region binding site was done using the Caver 1.0 software [36]. One of the most relevant tasks here is to appropriately select the tunnel starting point residues, which are typically placed within the active site. For this purpose, we consider the three fibrinogen binding sites with the highest D_g values and redundant tunnels were removed in order to set the starting point residues, namely: at *site 1*: N:48 N:49 N:50 N:51 P:19 P:20 P:21 P:22 O:81 O:83 O:84 O:85 O:86 Q:48 Q:49 Q:50 Q:51 S:19 S:20 S:21

S:22 R:81 R:83 R:84 R:85 R:86; at *site 2*: O:76 O:78 O:79 Q:54 Q:57 Q:60 Q:61 S:26 S:29 S:30 S:33 S:34; and at *site 3*: N:54 N:57 N:60 N:61 P:26 P:29 P:30 O:64 P:33 R:76 R:78 R:79. The algorithm implemented in this software allows to calculate the geometry of a given binding site and the target-residues composing the tunnel, as well as searching the tunnel pathways with the minimal radius and the lowest cost based on a Voronoi tessellation procedure [38]. In all the cases, the particular detection parameters applied here are minimum probe radius = 0.9 Å, shell depth = 4 Å, shell radius = 3 Å, clustering threshold = 3.5 Å, maximal distance = 3 Å, and desired radius = 5 Å.

Beta-blocker ligand-transport analysis. Afterwards, we carried out a ligand transport analysis (LTA) on the previously detected tunnels using the CaverDock software [54]. The LTA algorithm is able to generate the trajectories from the individual free binding energy (ΔG_{bind}) by performing an iterative docking of the beta-blockers (A or P) to every slice of the tunnel. In so doing, we employed the partially restrained AutoDock Vina scoring function [65] that considers two relevant parameters, *i.e.*: (i) the discretization delta that defines the distance between two slices centers of the tunnel, and (ii) the calculation mode that defines which ligand restraints will be set as rotation restriction coupled with backtracking to ensure a forward continuous movement. For that, the beta-blocker free binding energy trajectories were set as default parameters, which is strongly recommended in pharmacodynamics DDI studies [54, 65].

In addition, to evaluate the efficiency of the simultaneous transport of the beta-blockers A and P through the different tunnels, the biophysical throughput parameter (***T***) was estimated according to the **equation 5**. The throughput of a given pathway can adopt the form of the equation below when the ligand transport trajectory has continuous variation through the tunnel radius [54].

$$\boldsymbol{T} = e^{\int_0^L r(l)^{-n}\, dl} \quad \textbf{(5)}$$

In this **equation 5**, L represents the total path length with radius r, the function $r(l)$ defines the largest radius for absence of collisions or steric clashes along the beta-blocker drug-drug trajectories, and $n \in N$, though its default value is $n = 2$ for tunnels with relatively large width. Here, the "*cost*" is obtained by numerically integrated the exponent using the trapezoidal rule with a uniform grid, a minimum number of trapezoids (= 8), and a minimum grid distance (= 0.1 Å). As such, the ***T*** values range from 0 to 1 [54].

Analysis of 3D interaction diagrams. To evaluate the relevant beta-blockers DDI between A plus P with the fibrinogen E-region, 3D interaction diagrams were plotted and analyzed using the ezCADD software [44, 45]. This software determines the non-covalent intermolecular interactions in each receptor-ligand complex. Then, its utility tool named ezLigPlot automatically represents schematic 3D-interaction diagrams for the van der Waals hydrophobic, H-bond, cation-π, and π-π stacking interactions if any, along with the corresponding interatomic distances for all the docking complexes [44, 45].

Step 3: Calculation of energetic contributions for binding affinity. The total binding affinity for the beta-blockers drug-drug/fibrinogen E-region complexes was computed as the sum of the individual binding energies of the beta-blockers A and P (ΔG^{A}_{bind} and ΔG^{P}_{bind}) in the three best-ranked fibrinogen E-region binding sites (from now on, denoted as sites 1, 2 and 3). To this end, the empirical force-field parameters of the Autodock Vina scoring function were used to ascertain the effect of the different thermodynamic contributions on the total binding affinity [41, 42, 49, 55, 56]. That is, the individual binding energies are obtained by summing up the different contributions, as shown below for *e.g.* the beta-blockers (acebutolol (A) or propranolol (P) represented by the **equation 6**, and then added together to gather the total binding affinity (ΔG^{T}_{bind} ; **eq 7**).

$$\Delta G^{A}_{bind} = \Delta G^{A}_{Gauss1} + \Delta G^{A}_{Gauss2} + \Delta G^{A}_{repulsion} + \Delta G^{A}_{H\text{-}bond} + \Delta G^{A}_{hydrophobic} + \Delta G^{A}_{rot}$$

$$\Delta G^{T}_{bind} = \Delta G^{A}_{bind} + \Delta G^{P}_{bind}$$

(**equations 6** and **7**, respectively)

In **eq 6**, the steric interaction-based Vina scoring function is calculated using three terms, *i.e.*: two attractive Gaussian functions distance dependent terms (ΔG_{Gauss1} and ΔG_{Gauss2}), which characterize the London dispersion forces, and one repulsive parabolic interaction term ($\Delta G_{repulsion}$), which reproduces the canonical Lennard-Jones interaction shape of beta-blockers with fibrinogen [41, 42, 49, 55, 56, 66]. The $\Delta G_{H\text{-}bond}$ term accounts for the pair of donor and acceptor hydrogen-bonds. The fourth term $\Delta G_{hydrophobic}$ aims at evaluating the non-covalent van der Waals interactions [33-35]. Lastly, the ΔG_{rot} term holds for the number of rotatable bonds/torsions in the beta-blocker ligands A or P.

The total binding affinity is therefore used to quantify the DDI when both beta-blockers interact simultaneously at the same biophysical environment and crystallographic plane of the three best-ranked fibrinogen E-region binding sites. When $\Delta G^{T}_{\text{bind}}$ > 0 kcal/mol, it is classified as energetically unfavorable, pointing either extremely low or complete absence of affinity of the formed complexes, otherwise it is categorized like medium to high binding affinity. Additional details about the thermodynamic force field parameters of the Autodock Vina scoring function used in this work can be seen in Refs [33-35, 66]. An additional control simulation experiment was carried out to evaluate the maximum *per* atom-energy contribution to the individual binding affinity of the A and P separately in the three best-ranked fibrinogen E-region binding sites.

5.3.2 Depth profiles of fibrinogen target-residues. In order to evaluate the binding relevance of the beta-blocker DDI, we have performed a depth profile analysis of the target-residues (***i***) including the target-chains (***c***) making the tunnels [37, 39, 40, 67]. To this end, the residue depth is represented by the $D_{[c,i]}$ parameter, defined according to the general following equation in compact notation:

$$D_{[c,i]} = \prod_{c}\left[1-\prod_{i}(1-q_i^c)\right]\times S_{\max} \quad \textbf{(8)}$$

where q_i^c , q_i^cknown as the *depth quasi-sequence-order* descriptor, represents an adjusted binding probability-based weighted average derived from the distance matrix between the twenty standard amino-acids, and displays a linear correlation with the maximum solvent accessibility $S_{\max}$, *i.e.*, the percentage of predicting cavity waters to be displaced from the cited critical residues ***i*** in the main-chains ***c*** [37, 39, 40, 67]. Also, the q_i^c descriptor is a function of the *quasi-sequence-position-order* descriptor (J_i), so-called *conservation score*, which is related to the position and occurrence of the residue ***i*** in the sequence of the chain. More details about these descriptors will be discussed in the next section.

It is also important to highlight that the $D_{[c,i]}$ parameters for two different chains are independent from each other. Our hypothesis here is that the depth profile of target-residues (***i***) surrounding tunnels can significantly modulate the beta-blocker drug-drug interactions. Following this idea, in the present work, three categories were defined to evaluate the binding relevance of the fibrinogen target-chains, taking into account the

number of chains (***c***) composing the tunnels involved in the beta-blocker DDI, namely: ***c*** = 6 (highest relevance), ***c*** = 3-4 (medium relevance), and ***c*** = 1-2 (low relevance).

Step 4: DFT modeling of beta-blocker drug-drug binding systems. A DFT study of the configurations for the best-ranked beta-blocker drug-drug binding complexes was performed [68]. To this end, the Kohn-Sham equations were solved with numerical atomic orbitals using the SIESTA code [57, 69]. Specifically, the Kohn-Sham orbitals are expanded in a finite basis set of numerical pseudoatomic orbitals, the latter being described by double-zeta polarization (DZP) functions [70]. Correlation energies were set by the local density approximation (LDA), as proposed by Perdew and Zunger [71]. In fact, the LDA approach has been shown to be more suitable than the generalized gradient approximation (GGA) to study weakly interacting systems with the presence of π-π stacking interactions [71].

Improved Troullier-Martins pseudopotentials were used to describe the interactions between the core and the valence electrons [58-60], and for the molecular orbitals a localized DZP polarization basis set was implemented. For all the beta-blockers DDI systems, a cutoff value of 200 Ry was used to represent the charges density in the grid integration. The optimization of the systems' structure was performed by a conjugate gradient method; the beta-blocker atomic positions being fully relaxed until the remaining forces acting on them dropped below 0.05 eV/Å.

The DFT-binding energies (E_b) between the beta-blockers A and P were calculated according to the following **equation 9** [58-60]:

$$E_b = -(E[\mathrm{A}+\mathrm{P}]_{\mathrm{total}} - E[\mathrm{A}] - E[\mathrm{P}]) \quad \textbf{(9)}$$

where $E[\mathrm{A}+\mathrm{P}]_{\mathrm{total}}$ stands for the total DFT energy of the beta-blocker drug-drug system, and $E[\mathrm{A}]$ or $E[\mathrm{P}]$ for that of the isolated beta-blocker molecules A or P, respectively.

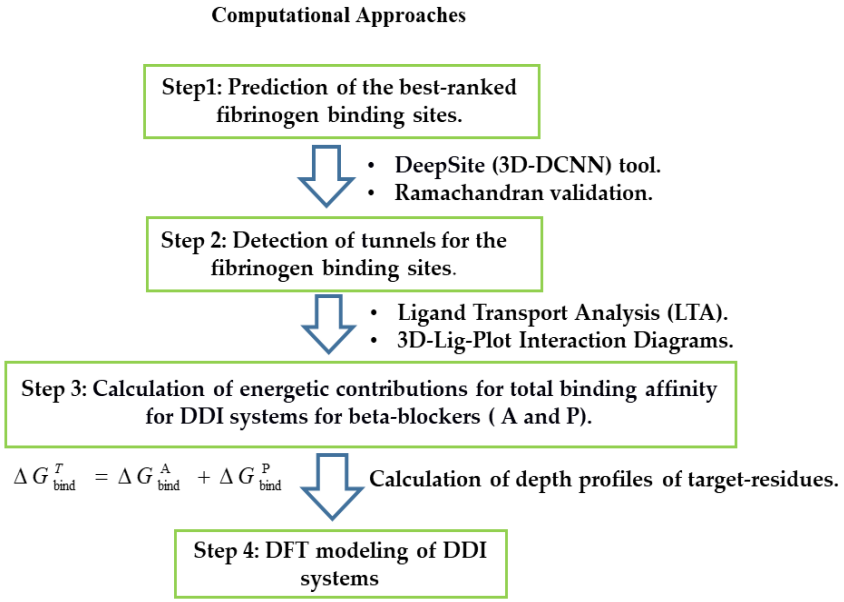


Figure 13. General flowchart of the computational methods applied. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

References

[1] J. Niu, R.M. Straubinger, D.E. Mager, Pharmacodynamic Drug–Drug Interactions, Clinical Pharmacology & Therapeutics 105(6) (2019) 1395-1406.

[2] A.D. Boran, R. Iyengar, Systems approaches to polypharmacology and drug discovery, Current opinion in drug discovery & development 13(3) (2010) 297.

[3] J. Mestres, E. Gregori-Puigjané, Conciliating binding efficiency and polypharmacology, Trends in pharmacological sciences 30(9) (2009) 470-474.

[4] J.-U. Peters, Polypharmacology foe or friend? Journal of medicinal chemistry 56(22) (2013) 8955-8971.

[5] Y. Hu, J. Bajorath, Compound promiscuity: what can we learn from current data? Drug discovery today 18(13-14) (2013) 644-650.

[6] Z. Simon, Á. Peragovics, M. Vigh-Smeller, G. Csukly, L. Tombor, Z. Yang, G. Zahoránszky-Kőhalmi, L. Végner, B. Jelinek, P. Hari, Drug effect prediction by polypharmacology-based interaction profiling, Journal of chemical information and modeling 52(1) (2011) 134-145.

[7] R. Chaudhari, Z. Tan, B. Huang, S. Zhang, Computational polypharmacology: a new paradigm for drug discovery, Expert opinion on drug discovery 12(3) (2017) 279-291.
[8] J.M. Ruso, A. González-Pérez, G. Prieto, F. Sarmiento, A volumetric study of two related amphiphilic beta-blockers as a function of temperature and electrolyte concentration, Colloids and Surfaces B: Biointerfaces 33(3-4) (2004) 165-175.
[9] G. Johnsson, C.G. Regardh, Clinical pharmacokinetics of beta-adrenoreceptor blocking drugs, Clinical pharmacokinetics 1(4) (1976) 233-63.
[10] U. Borchard, Pharmacokinetics of beta-adrenoceptor blocking agents: clinical significance of hepatic and/or renal clearance, Clinical physiology and biochemistry 8 Suppl 2 (1990) 28-34.
[11] E. Lounkine, M.J. Keiser, S. Whitebread, D. Mikhailov, J. Hamon, J.L. Jenkins, P. Lavan, E. Weber, A.K. Doak, S. Côté, Large-scale prediction and testing of drug activity on side-effect targets, Nature 486(7403) (2012) 361.
[12] M. Zitnik, M. Agrawal, J. Leskovec, Modeling polypharmacy side effects with graph convolutional networks, Bioinformatics 34(13) (2018) i457-i466.
[13] A. Moya-García, T. Adeyelu, F.A. Kruger, N.L. Dawson, J.G. Lees, J.P. Overington, C. Orengo, J.A. Ranea, Structural and functional view of polypharmacology, Scientific reports 7(1) (2017) 10102.
[14] A. Zimmer, I. Katzir, E. Dekel, A.E. Mayo, U. Alon, Prediction of multidimensional drug dose responses based on measurements of drug pairs, Proceedings of the National Academy of Sciences 113(37) (2016) 10442-10447.
[15] S. Loewe, Effect of combinations: mathematical basis of problem, Arch. Exp. Pathol. Pharmakol. 114 (1926) 313-326.
[16] J. Lehár, G.R. Zimmermann, A.S. Krueger, R.A. Molnar, J.T. Ledell, A.M. Heilbut, G.F. Short, L.C. Giusti, G.P. Nolan, O.A. Magid, Chemical combination effects predict connectivity in biological systems, Molecular systems biology 3(1) (2007).
[17] J.Y. Ryu, H.U. Kim, S.Y. Lee, Deep learning improves prediction of drug–drug and drug–food interactions, Proceedings of the National Academy of Sciences 115(18) (2018) E4304-E4311.
[18] J. Owens, Target validation: Determining druggability, Nature Reviews Drug Discovery 6(3) (2007) 187.
[19] A.C. Cheng, R.G. Coleman, K.T. Smyth, Q. Cao, P. Soulard, D.R. Caffrey, A.C. Salzberg, E.S. Huang, Structure-based maximal affinity model predicts small-molecule druggability, Nature biotechnology 25(1) (2007) 71.

[20] A. Bakan, N. Nevins, A.S. Lakdawala, I. Bahar, Druggability assessment of allosteric proteins by dynamics simulations in the presence of probe molecules, Journal of chemical theory and computation 8(7) (2012) 2435-2447.

[21] P. Pasznik, E. Rutkowska, S. Niewieczerzal, J. Cielecka-Piontek, D. Latek, Potential off-target effects of beta-blockers on gut hormone receptors: In silico study including GUT-DOCK—A web service for small-molecule docking, PloS one 14(1) (2019) e0210705.

[22] N. Hassan, J. Maldonado-Valderrama, A.P. Gunning, V.J. Morris, J.M. Ruso, Surface characterization and AFM Imaging of mixed fibrinogen− surfactant films, The Journal of Physical Chemistry B 115(19) (2011) 6304-6311.

[23] N. Hassan, J.M. Ruso, P. Somasundaran, Mechanisms of fibrinogen–acebutolol interactions: Insights from DSC, CD and LS, Colloids and Surfaces B: Biointerfaces 82(2) (2011) 581-587.

[24] G. Pucci, M.G. Ranalli, F. Battista, G. Schillaci, Effects of β-blockers with and without vasodilating properties on central blood pressure: systematic review and meta-analysis of randomized trials in hypertension, Hypertension 67(2) (2016) 316-324.

[25] A. Bratek-Skicki, P. Żeliszewska, J.M. Ruso, Fibrinogen: a journey into biotechnology, Soft matter 12(42) (2016) 8639-8653.

[26] J. Madrazo, J.H. Brown, S. Litvinovich, R. Dominguez, S. Yakovlev, L. Medved, C. Cohen, Crystal structure of the central region of bovine fibrinogen (E5 fragment) at 1.4-Å resolution, Proceedings of the National Academy of Sciences 98(21) (2001) 11967-11972.

[27] J.M. Kollman, L. Pandi, M.R. Sawaya, M. Riley, R.F. Doolittle, Crystal structure of human fibrinogen, Biochemistry 48(18) (2009) 3877-3886.

[28] I. Pechik, S. Yakovlev, M.W. Mosesson, G.L. Gilliland, L. Medved, Structural basis for sequential cleavage of fibrinopeptides upon fibrin assembly, Biochemistry 45(11) (2006) 3588-3597.

[29] J.W. Weisel, J.G. Phillips, C. Cohen, The structure of fibrinogen and fibrin: II. Architecture of the fibrin clot, Annals of the New York Academy of Sciences 408 (1983) 367-379.

[30] J.H. Brown, N. Volkmann, G. Jun, A.H. Henschen-Edman, C. Cohen, The crystal structure of modified bovine fibrinogen, Proceedings of the National Academy of Sciences 97(1) (2000) 85-90.

[31] N. Hassan, L.R. Barbosa, R. Itri, J.M. Ruso, Fibrinogen stability under surfactant interaction, Journal of colloid and interface science 362(1) (2011) 118-126.
[32] J. Jiménez, S. Doerr, G. Martínez-Rosell, A.S. Rose, G. De Fabritiis, DeepSite: protein-binding site predictor using 3D-convolutional neural networks, Bioinformatics 33(19) (2017) 3036-3042.
[33] M. Rarey, B. Kramer, T. Lengauer, Multiple automatic base selection: Protein–ligand docking based on incremental construction without manual intervention, Journal of computer-aided molecular design 11(4) (1997) 369-384.
[34] D. Seeliger, B.L. de Groot, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, Journal of computer-aided molecular design 24(5) (2010) 417-422.
[35] B.K. Shoichet, Virtual screening of chemical libraries, Nature 432(7019) (2004) 862.
[36] J. Stourac, O. Vavra, P. Kokkonen, J. Filipovic, G. Pinto, J. Brezovsky, J. Damborsky, D. Bednar, Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport, Nucleic acids research (2019).
[37] V. Le Guilloux, P. Schmidtke, P. Tuffery, Fpocket: an open source platform for ligand pocket detection, BMC bioinformatics 10(1) (2009) 168.
[38] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallographica Section D: Biological Crystallography 66(1) (2010) 12-21.
[39] K.P. Tan, T.B. Nguyen, S. Patel, R. Varadarajan, M.S. Madhusudhan, Depth: a web server to compute depth, cavity sizes, detect potential small-molecule ligand-binding cavities and predict the pKa of ionizable residues in proteins, Nucleic acids research 41(W1) (2013) W314-W321.
[40] K.P. Tan, R. Varadarajan, M.S. Madhusudhan, DEPTH: a web server to compute depth and predict small-molecule binding cavities in proteins, Nucleic acids research 39(suppl_2) (2011) W242-W248.
[41] S. Forli, R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, A.J. Olson, Computational protein–ligand docking and virtual drug screening with the AutoDock suite, Nature protocols 11(5) (2016) 905.
[42] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, Journal of computational chemistry 19(14) (1998) 1639-1662.

[43] M. González-Durruthy, G. Scanavachi, R. Rial, Z. Liu, M.N.D. Cordeiro, R. Itri, J.M. Ruso, Structural and energetic evolution of fibrinogen toward to the betablocker interactions, International journal of biological macromolecules 137 (2019) 405-419.
[44] A. Tao, Y. Huang, Y. Shinohara, M.L. Caylor, S. Pashikanti, D. Xu, ezCADD: A Rapid 2D/3D Visualization-Enabled Web Modeling Environment for Democratizing Computer-Aided Drug Design, Journal of chemical information and modeling 59(1) (2018) 18-24.
[45] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, ACS Publications, 2011.
[46] N.E. Newell, Mapping side chain interactions at protein helix termini, BMC bioinformatics 16(1) (2015) 231.
[47] A. Bachmann, D. Wildemann, F. Praetorius, G. Fischer, T. Kiefhaber, Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction, Proceedings of the National Academy of Sciences 108(10) (2011) 3952-3957.
[48] M.I. Stefan, S.J. Edelstein, N. Le Novère, Computing phenomenologic Adair-Klotz constants from microscopic MWC parameters, BMC systems biology 3(1) (2009) 68.
[49] J. Haiech, Y. Gendrault, M.-C. Kilhoffer, R. Ranjeva, M. Madec, C. Lallement, A general framework improving teaching ligand binding to a macromolecule, Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1843(10) (2014) 2348-2355.
[50] J.E. Fletcher, A.A. Spector, J.D. Ashbrook, Analysis of macromolecule-ligand binding by determination of stepwise equilibrium constants, Biochemistry 9(23) (1970) 4580-4587.
[51] Z. Konkoli, Safe uses of Hill's model: an exact comparison with the Adair-Klotz model, Theoretical Biology and Medical Modelling 8(1) (2011) 10.
[52] V.M. de Menezes, E. Michelon, J. Rossato, I. Zanella, S.B. Fagan, Carbon nanostructures interacting with vitamins A, B3 and C: ab initio simulations, Journal of biomedical nanotechnology 8(2) (2012) 345-349.
[53] I. Sánchez-Linares, H. Pérez-Sánchez, J.M. Cecilia, J.M. García, High-throughput parallel blind virtual screening using BINDSURF, BMC bioinformatics 13(14) (2012) S13.
[54] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, CAVER 3.0: a tool for the analysis of transport

pathways in dynamic protein structures, PLoS computational biology 8(10) (2012) e1002708.

[55] Z.-R. Xie, M.-J. Hwang. An interaction-motif-based scoring function for protein-ligand docking, BMC bioinformatics 11(1) (2010) 298.

[56] H.A. Feldman, Mathematical theory of complex ligand-binding systems at equilibrium: some methods for parameter fitting, Analytical biochemistry 48(2) (1972) 317-338.

[57] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, The SIESTA method for *ab initio* order-N materials simulation, Journal of Physics: Condensed Matter 14(11) (2002) 2745.

[58] Y. Wang, Z. Xu, Y.N. Moe, On the performance of local density approximation in describing the adsorption of electron donating/accepting molecules on graphene, Chemical physics 406 (2012) 78-85.

[59] M. Arrigoni, G.K. Madsen, Comparing the performance of LDA and GGA functionals in predicting the lattice thermal conductivity of III-V semiconductor materials in the zinc blende structure: The cases of AlAs and BAs, Computational Materials Science 156 (2019) 354-360.

[60] N. Troullier, J.L. Martins, Efficient pseudopotentials for plane-wave calculations, Physical review B 43(3) (1991) 1993.

[61] J. Westbrook, Z. Feng, L. Chen, H. Yang, H.M. Berman, The protein data bank and structural genomics, Nucleic acids research 31(1) (2003) 489-491.

[62] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, Journal of computational chemistry 31(2) (2010) 455-461.

[63] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, PubChem substance and compound databases, Nucleic acids research 44(D1) (2015) D1202-D1213.

[64] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, Journal of cheminformatics 4(1) (2012) 17.

[65] W.P. Feinstein, M. Brylinski. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets, Journal of cheminformatics 7(1) (2015) 18.

[66] R. Quiroga, M.A. Villarreal, Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening, PLOS ONE 11(5) (2016) e0155183.
[67] M.A. Marti-Renom, M. Madhusudhan, A. Sali, Alignment of protein sequences by their profiles, Protein Science 13(4) (2004) 1071-1087.
[68] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Physical review 136(3B) (1964) B864.
[69] W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects, Physical review 140(4A) (1965) A1133.
[70] O.F. Sankey, D.J. Niklewski, Ab initio multicenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems, Physical Review B 40(6) (1989) 3979.
[71] J.P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Physical Review B 23(10) (1981) 5048.

Supporting Information

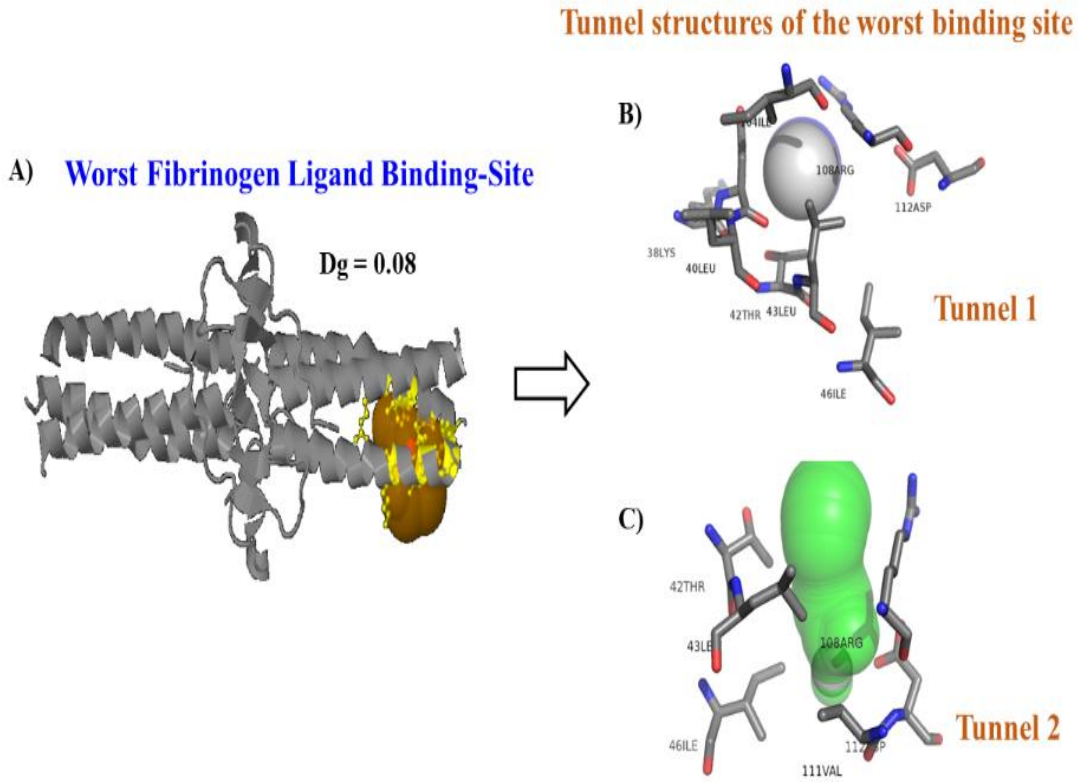


Figure S1. **A)** Prediction of the worst catalytic binding site of fibrinogen E-region based on druggability degree (Dg = 0.08). **B)** Tunnel detection of the worst binding site of fibrinogen E-region like tunnel 1 (gray) and tunnel 2 (green) with the surrounding catalytic residues. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

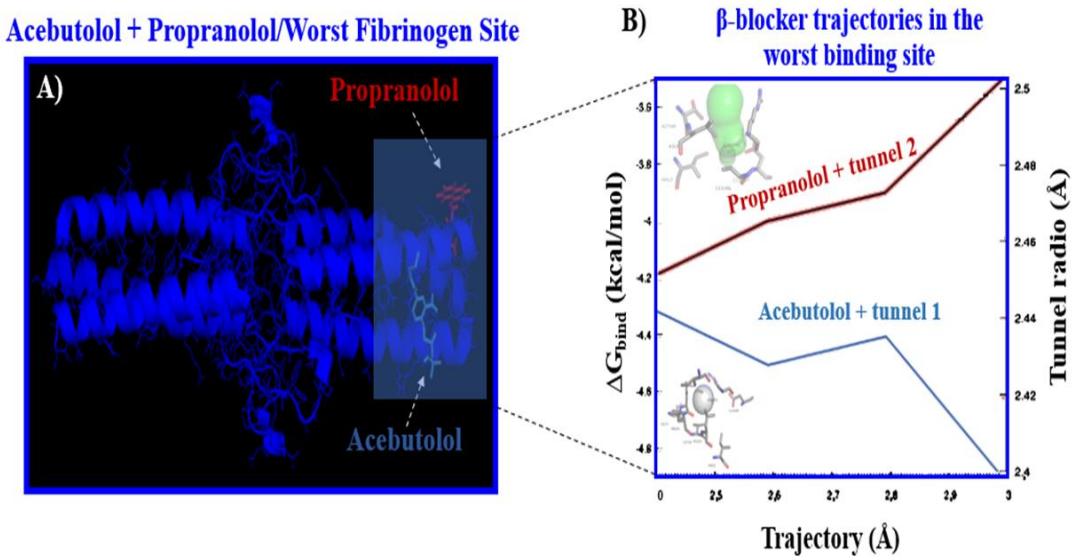


Figure S2. **A)** Cartoon representation of β-blocker interactions like acebutolol + tunnel 1 (blue) and propranolol + tunnel 2 (red) from the worst fibrinogen binding site (Dg = 0.08). **B)** Binding profiles of β-blocker trajectories showing the total absence of β-blocker drug-drug interactions based on the non-interception between β-blocker trajectories. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

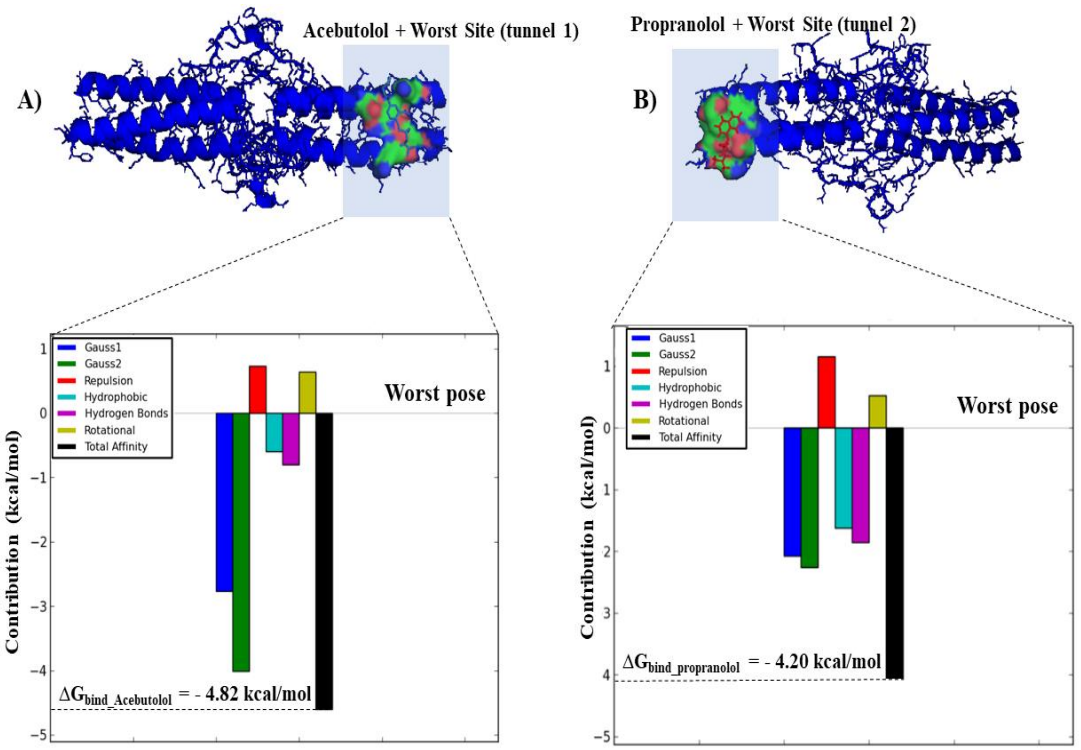


Figure S3. Graphical breakdown of the different binding energy contributions (ΔG_{Gauss1}, ΔG_{Gauss2}, $\Delta G_{repulsion}$, $\Delta G_{H\text{-}bond}$, $\Delta G_{hydrophobic}$, and $\Delta G_{rotational}$) to the individual binding affinity (ΔG_{bind}, kcal/mol) of the β-blockers like **A)** acebutolol/worst site_tunnel 1 ($\Delta G_{bind\_Acebutolol}$ = - 4.82 kcal/mol) and **B)** propranolol/worst site_tunnel 2 ($\Delta G_{bind\_Propranolol}$ = - 4.20 kcal/mol). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

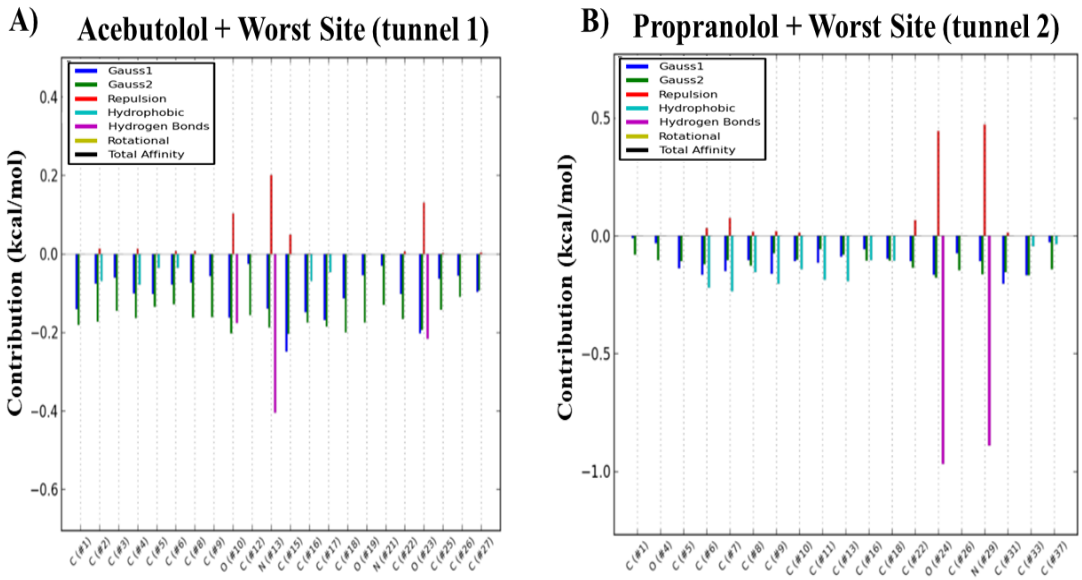


Figure S4. Representation of the per atom energy contributions (kcal/mol) to the individual binding affinity (ΔG_{bind}) of the acebutolol ($\Delta G_{bind\_Acebutolol}$) and propranolol ($\Delta G_{bind\_Propranolol}$) in the worst fibrinogen binding site (Dg = 0.08). **A)** acebutolol + worst site (tunnel 1) and **B)** propranolol + worst site (tunnel 2). Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

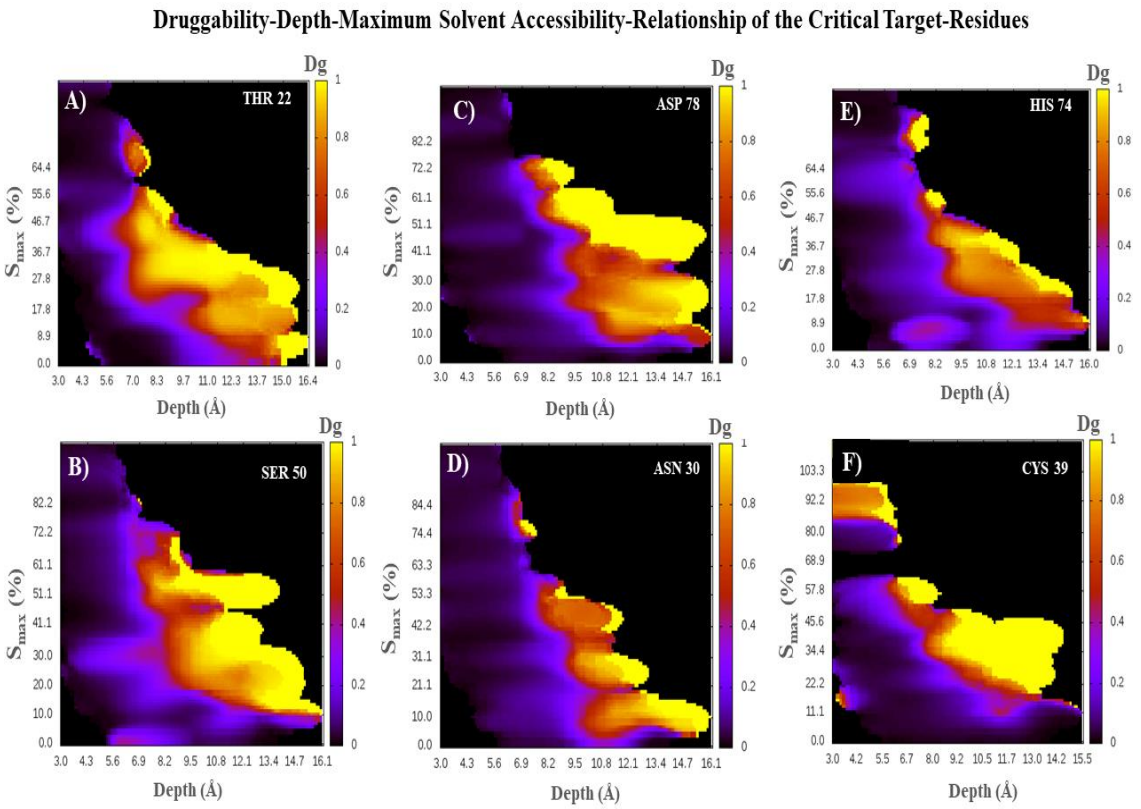


Figure S5. Druggability-depth-maximum solvent accessibility-relationship of the critical target-residues belonging to fibrinogen E-region binding sites (site 1, site 2, and site 3). **A)** TRH22, **B)** SER50, **C)** ASP78, **D)** ASN30, **E)** HIS74, **F)** CYS39. Source MDPI article, please visit: https://doi:10.3390/molecules25225425 (accessed on 19 november 2020)

Chapter 6.

Michael González-Durruthy, Ramón Rial, M. Natália D.S. Cordeiro, Zhen Liu, Juan M. Ruso. Exploring the conformational binding mechanism of fibrinogen induced by interactions with penicillin β-lactam antibiotic drugs. *Journal of Molecular Liquids*, 324, 114667, 2021. doi.org/10.1016/j.molliq.2020.114667.

Journal Impact Factor (JIF) (2021): 6.165.

CiteScore (2021): 6.9

Author Contribution M.G.-D.; Conceptualization, methodology-based on molecular docking, flexibility properties evaluations, building perturbation response scanning maps, and writing—original draft preparation. *Corresponding author

Journal authorization:

Chapter 6.

Exploring the conformational binding mechanism of fibrinogen induced by interactions with penicillin β-lactam antibiotic drugs.

Abstract

Herein, we present an integrated computational and experimental study to tackle the interactions between recognized β-lactam antibiotics (cloxacillin and dicloxacillin) with the fibrinogen blood plasma protein. For this purpose, molecular docking simulation with elastic network based on collective low-frequency normal modes and perturbation response scanning maps were proposed to evaluate the conformational binding mechanism of fibrinogen under the unbound and bound states with the cited β-lactam antibiotics. Aiming to theoretically explore the hidden biochemical mechanisms and structural attributes leading failures in therapy success with β-lactam antibiotics. The computational results pointing that despite these conformational differences, both antibiotics exhibit very similar affinity-based free energies of binding as FEB (cloxacillin/E-region) = - 8.7 kcal/mol and FEB (dicloxacillin/E-region) = -7.7 kcal/mol. We theoretically suggest that the semi-synthetic incorporation of an additional halogen CL-atom in the dicloxacillin, respect to cloxacillin molecule, and its relative docking-pose orientation in the fibrinogen E-region could significantly reduce the appearance of potential fibrinolytic off-target effects usually associated to parenterally administered β-lactam antibiotics. Besides, the performed interactions diagrams revealed that the most relevant antibiotic binding interactions with the fibrinogen E-region (pocket 1) are mainly based on hydrophobic (C⋯C)-backbone-side-chain non-covalent interactions, acceptor/donor interactions with critical regulatory E-region residues SER50:Q > SER50:N associated to allosteric modulation based long-distance-based perturbations (dicloxacillin >> cloxacillin) in the E-region (Q-chain > N-chain) with remarkable conformational rigidification by decreasing the intrinsic collectivity, and leading different pattern of perturbations as allosteric signal propagation in the intrinsic conformational dynamics under bound state from both β-lactam antibiotics. Overall the obtained theoretical evidences fit well with previous experimental ones obtained by our group by using isothermal titration calorimetry (ITC) and quantitative differential scanning calorimetry (DSC) (refer to https://doi.org/10.1016/j.molliq.2020.114667). Allowing to

corroborate and validate the proposed computational approaches ones. Finally, the obtained results open new perspectives for the "*de novo rational drug-design"* of new derivatives of β-lactam antibiotics with high pharmacodynamic selectivity/specificity to avoid side-effects toward to achieve optimal benefit/risk rates beyond the antibiotic drug resistance phenomena, favoring the implementation of rigorous criteria for a more personalized antibiotic therapy.

6.1 Introduction

Despite the enormous attention that for so many years have enjoyed by the scientific community protein-ligand interactions, far from diminishing, has increased the interest and fascination that this topic exercise upon researchers. Many are the causes that explain such an attraction: multidisciplinary approach, constant discovery of new proteins and ligands, the recent emergence of COVID-19 with the consequent demand for new vaccines and medicines[1], the development of artificial intelligence methods that are nurtured by large databases[2], the excellent combination of thriving tools both experimental and computing that this field more than others need to coexist [3] and the scientific challenge that the complexity of these systems (dynamic, fluctuations, degrees of freedom, crowding, etc.) assumes[4].

Penicillin β-lactam antibiotics are recognized as one of the most common wide-spectrum antibiotics used for different infections. However, the therapy success with penicillin β-lactam antibiotics remains as unsolved problem since β-lactamase resistance due to the inappropriate use, is disseminating rapidly among pathogenic bacteria. Besides, considering the occurrence of potential hematotoxicity (off-target fibrinolytic interactions) as hemolytic anemia caused by intravascular hemolysis, and blood clotting disorders-mediated fibrinolysis (i.e., by interactions with the fibrinogen molecule) which are recognized as one of the most severe adverse reactions induced by the penicillin β-lactam antibiotics when are parenterally administered [5]. To overcome the potential hematotoxicity of the β-lactam antibiotics therapeutic strategies should consider the rational prescription, substitution and/or combination [6].

In pursuing these aims, we have studied the interactions and complexation of two penicillin β-lactam antibiotics, cloxacillin and dicloxacillin, with the fibrinogen serum protein. We have attempted to correlate the effects of chemical substitution on molecular and thermodynamic properties of protein in the bulk with their effects on protein functional properties. The penicillin drugs selected for study form an interesting series of

molecules in which the only variation in the molecular structure is the number and nature of the substituents on the aromatic ring: an additional chlorine atom on the phenyl ring of dicloxacillin. This selection provides an interesting opportunity to investigate the relationships of the molecular structure on the intermolecular interactions[7-9]. In previous studies we investigated the relationships between molecular architecture and the physicochemical properties of structures that are formed by these molecules in aqueous solution. Static light scattering and NMR studies have shown that both drugs form small aggregates from a given concentration in aqueous solution, but there was evidence of a second aggregation in solutions of dicloxacillin (but not cloxacillin)[10]. The stability of these aggregates has not yet been quantified sufficiently, but it has highlighted its enormous importance on the bacterial activity [11] and protein interactions, where evidence has been presented for the haptenation of penicillin by conjugation with serum proteins which results in the formation of IgE antibodies [12, 13]. From the structural and functional point of view the human fibrinogen is a soluble glycoprotein composed by disulfide-linked dimer of three nonidentical polypeptide chains, Aα, Bβ, and γ. The NH_2 terminal portions of the six chains are linked together to the central region (E) of the molecule by 11 disulfide bonds forming a small globular domain, the so-called disulfide knot, in the center. The C termini of each of the three chains end in globular domains, those of the Bβ, and γ chains are located at the ends of the molecule. The COOH-terminal portion of each fibrinogen Aα chain forms a compact αC-domain attached to the bulk of the molecule with a flexible αC-connector. In addition, fibrinogen shows a unique characteristic in its folding[14]. According to the current view, in fibrinogen, two αC-domains interact intramolecularly with each other and with the central region of the molecule, while in fibrin, they switch to an intermolecular interaction to form αC-polymers. The central region (E-region) is responsible for the fibrin polymerization during the clotting process[15]. Fibrinogen circulates in the plasma with a molecular weight of ~340 kDa depending of Aα and Bβ chains and the γ-γ and α-α crosslinking chains content. It has a shape similar to a rod with dimensions of 9 × 47.5 × 6 nm and present a negative net charge with isoelectric pH=5.8 in physiological conditions[16]. The normal concentration of fibrinogen is around 150-400 mg/dL in the blood plasma. Levels noticeably below or above this range are associated with pathological conditions like bleeding and/or thrombosis [17]. Crystallographic analysis of the molecule has been hindered for a long time by its complex and flexible structure. Finally, a planar sigmoidal structure was proposed: the axis of the α-helical coiled-coil rod adopts a sigmoidal shape

that lies nearly in a plane, as the curvature is an intrinsic feature of fibrinogen's coiled coil [18].

In order to achieve these objectives and based on all the above, we have approached this task with a combination computational methods. In this context, computational methods like molecular docking approaches has become an increasingly important and recognized tool in rational drug-design[19] The molecular docking simulation can be applied to model the interaction between a small molecule and a protein receptor at the atomic level, allowing to efficiently elucidate keys biochemical processes [20]. Despite the existence of great advances in computational pharmacology are still ignored how different binding-modes can modulate the pharmacological response. The molecular docking results generally suggest that the native binding-mode corresponds to a low free energy structure but not necessarily with the lowest values. In this regard, one of the main challenges, is to implement computational algorithms that allow to discriminate correctly the influence of a given conformational binding-modes of the ligand on the interaction energy (FEB values), which are usually, very close to each other[21]. Because, in most cases the different ligand-conformations obtained after the molecular docking simulations occupy the same biophysical environment in the protein binding-site. For this instance, the local perturbation response scanning maps (LPRS map images) could be considered as an efficient approach to identify differences-based penicillin-conformational binding-modes with fibrinogen protein[22, 23]. The LPRS maps allows the determination of the influence/sensitivity that each binding residue has on/to every other residue under ligand interaction, by performing a map-based matrix of interaction that works with an anisotropic network model (ANM), where the nodes refer to individual residues[24, 25].

6.2 Materials and Methods

6.2.1 Computational modeling-based molecular docking.

To study the binding properties between the penicillin β-lactam antibiotics and fibrinogen protein a computational mechanistic study based on molecular docking simulation was performed. Towards such end, first we prepared the protein receptor file (*i.e.,* fibrinogen), which was withdrawn from the *RCSB Protein Data Bank* (PDB) X-ray structures, *i.e.* with *PDB ID*: 1JY2 [26]. Afterwards, the obtained fibrinogen E-region structure was prepared using an online *ProteinPrepare* tool [27]. The preparation process includes the titration of the protonation states using *PROPKA 3.1 tool*, addition of missing atoms, removal of co-crystallized ligands and overall optimization of the H-network using

PDB2PQR 2.1. Following this step, the penicillin β-lactam antibiotic ligands (cloxacillin and dicloxacillin) were obtained from the *Pubchem Data Base Chemical Structure Search* as cloxacillin (PubChem CID: 6098; MF: 435.9 g/mol) and dicloxacillin (PubChem CID: 18381; MF: 470.3 g/mol) [16]. The drugs geometry optimization was carried out by using the MOPAC extension based on NDDO approximation [28]. Then, to study the β-lactam antibiotic docking mechanisms with fibrinogen, we used the Autodock Vina scoring function developed by Trott *et al*.[29, 30] to obtain the free energy of binging (FEB) based on X-Score function which approximates the standard chemical potentials. Before the docking modeling, the fibrinogen binding-sites were predicted through the ezPocket with *fconv* for binding-site detection [31, 32]. This step was carried out as delimiting the access to fibrinogen-cavities, like van der Waals surfaces that are likely to bind to the ligands. To this end, the *fconv* analysis uses Delaunay triangulation with weighted points to detect all the relevant binding pockets. Herein, the volumetric map of the fibrinogen binding-site is generated together with the Cartesian *XYZ*-coordinates for the docking box simulations [33], *i.e.*: grid box size with dimensions of X= 64Å, Y= 52Å, Z= 52Å and grid box center X= 13.1Å, Y= -2.0Å, Z= 12.5Å. For this purpose, a docking accuracy was fixed at 100 and after that, the best conformations or penicillin binding-poses were selected [34]. The docking affinity (FEB values) were classified like energetically-unfavorable when the FEB of fibrinogen-penicillin complexes $\geq$ 0 kcal/mol, therefore indicating either extremely low or complete absence of affinity; otherwise the fibrinogen-penicillin docking complexes were classified like medium to high docking affinity [29].

6.2.2 Performing 2D-3D Lig-Plot interaction diagrams.

This algorithm was applied to evaluate the influence in the FEB values for the different contributions of the β-lactam antibiotics binding-poses interacting with the fibrinogen. To specify the key relevant intermolecular interactions 3D Lig-Plot diagrams were analyzed. To tackle this objective, the ezLigPlot software was used. This software determines the non-covalent intermolecular interactions present in a given protein-ligand complex and it builds automatically a 3D-interaction diagram that includes hydrophobic, H-bond, cation-Π, and Π-Π stacking interactions along with its corresponding interatomic distances (d_{ij}) for each binding-poses from the obtained docking complexes [34].

6.2.3 Building perturbation response scanning maps

This computational approach evaluates the degree of change (*i.e.*, interatomic distances perturbation between residue fluctuations) induced by the ligands in the residues network of fibrinogen by describing the interaction potential (*V*) from the receptor-ligand complex (R-L or fibrinogen-β-lactam antibiotics) as an anisotropic network model (ANM) of interacting springs based on the Elastic Network Model Theory (ENM) [23, 35-37], where each spring connects a pair of C(α)-nodes (i-j) corresponding to the ligands (*i.e.*, (i)-cloxacillin or (i)-dicloxacillin) and the receptor (j-fibrinogen) are placed within a given distance cutoff of 7 Å. Herein, we use the ANM models applying the concepts on rotation-translations blocks approximation (RTB model) to explore the collective motions in low-frequency normal modes. Where the target protein (fibrinogen) is splitted into n-(b)-blocks, a block being made of one or of a few consecutive residues that belong to the different polypeptide chains of the fibrinogen protein that could potentially interact with the penicillin β-lactam antibiotic ligands (cloxacillin and dicloxacillin). Then, the ANM-RTB model was implemented to evaluate the potential perturbation responses-induced by the β-lactam antibiotics on fibrinogen binding-site network of residues using local perturbation-response scanning maps (LPRS maps) to measure the fluctuations of the displacements from equilibrium between the unperturbed (unbound fibrinogen) and perturbed state (bound fibrinogen under the β-lactam antibiotic interactions) following **equations 1** and **2**:

$$V_R = \sum_{(i<j)\in N(R)} \frac{1}{2}\gamma_R \left(d_{ij} - d_{ij}^0\right)^2 \quad \textbf{(1)}$$

$$V'_{R-L} = \sum_{(i<j)\in N(R)\cup N(L)} \frac{1}{2}\gamma_R \left(d_{ij} - d_{ij}^0\right)^2 + \sum_{(i,j)\in N(R,L)} \frac{1}{2}\gamma_L \left(d_{ij} - d_{ij}^0\right)^2 \quad \textbf{(2)}$$

Where V describes the interaction potential for a protein receptor (*i.e.*, V_R: fibrinogen in the unperturbed like unbound state representing the intrinsic conformational dynamic in the absence of the penicillin β-lactam antibiotic ligands), and the bound state representing the interaction-based perturbations (*i.e.*, $V_{R\text{-}L}$: fibrinogen/cloxacillin and fibrinogen/dicloxacillin as generated docking complexes) modeled by a quasi-harmonic Hookean springs [38]. For this instance, the force constant (γ) of the spring connecting C(α)-atoms *i* and *j* from the receptor (interacting residues) and the ligand are taken as γ_R and γ_L respectively. See **equation 3**:

$$\gamma = exp\left[\left(-1/2\right)\left(\left(d_{ij}^{0}/d_{cutoff}\right)^{2}\right)\right] \quad (3)$$

The cutoff for the fibrinogen $\gamma_R = 1$, d_{ij}^0 is the distance between i,j-C(α)-atoms of the β-lactam antibiotic as ligand (*L*) and fibrinogen receptor (*R*), in the equilibrium condition indicated by the superscript "0", and the $d_{cutoff} = 7$ Å. More details will be explained in the next section.

6.3 Results and discussion

6.3.1 Fibrinogen structural characterization and binding site prediction

The accurate identification of druggable binding pockets constitutes an essential step for structure-based drug design antibiotics considering the long list of biochemical mechanisms associated to antibiotic-resistance phenomena[39]. Which is widely recognized as a fundamental public health threat affecting humans worldwide. In this context, the first step of molecular docking approach is the correct prediction/identification of a feasible binding pockets with an appropriate crystallographic structural characterization of their properties toward to ensure quality of modeling data as ligand selectivity and specificity in order to avoid adverse effects[27]. In the present study we address the prediction of druggable fibrinogen binding sites considering the fibrinogen whole structure (E-region, D-region 1 and D-region 2) in order to determine the highest concentration and patterns of distribution of active binding sites throughout the fibrinogen structure. Following this idea, the binding-sites prediction of fibrinogen was carried out by applying fpocket with ezPocket software which use Voronoi tessellation algorithm to detect/scanning concave surfaces and tiny sub-cavities (tunnels) collecting information-based on the crystallographic descriptors related to pocket topology and geometry to accommodate the ligands (cloxacillin and dicloxacillin) with favorable interactions. See **Figure 1.**

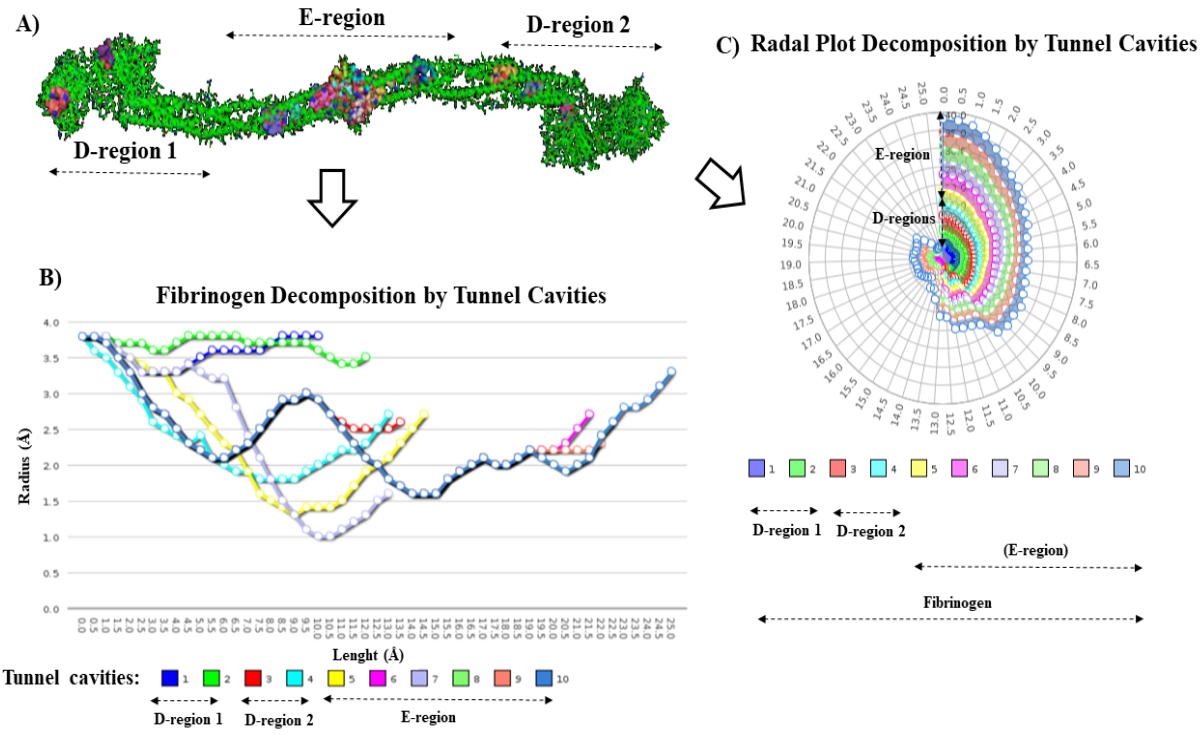


Figure 1. A) Representation of general crystallographic structure of fibrinogen protein by relevant regions formed by two quasi-symmetric C-terminal portions (D-regions 1 and 2) and the central part (E-region thrombin binding-domain) highlighting the binding sites formed by small cluster of hydrophobic cavities represented as colored van der Waals surfaces heterogeneously distributed across the overall structure. B) Graphical representation of length (Å) vs. radius (Å) obtained for the fibrinogen tunnel sub-cavities (10 entangled tunnels) belonging binding sites for each fibrinogen regions (E-region and the two quasi-symmetric D-regions 1 and 2). C) Radial plot decomposition of the tunnel sub-cavities of the fibrinogen binding sites showing the overlaping within the each fibrinogen region. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

For this instance, the fpocket method calculates the geometric center of the pocket and considering if the relative position of the pocket center is within 4 Å from any atom of the potential ligand to be considered as correctly predicted or identified. The fpocket binding cavity detection method can generates Cartesian coordinates-like van der Waals volumetric maps that allows setting the box-simulation used for docking experiments [32]. The obtained results show that the fibrinogen protein has multiple cavities more densely located in the N-terminal central nodule E-region which correspond to the

thrombin binding-domain which has critical importance for the blood coagulation process (fibrin polymerization). By the other hand, the regions composed by the two quasi-symmetric C-terminal portions showed the smallest quantity of binding sites (D-region 2 > D-region 1). In addition, we perform a detection of tiny sub-cavities (or tunnels) composing fibrinogen binding sites for each region [17, 40]. For this instance, a total of 10 relevant tunnels were identify, 6 interconnected tunnels composing the E-region binding sites, and 2 tunnels for each D-region binding sites. Herein, was clearly shown that the number of tunnel cavities forming the E-region binding sites are greater than the tunnel cavities composing from both D-regions. Furthermore, the length and radius of the tunnels composing the E-region binding sites are also significantly higher than the D-region sub-cavity tunnels as showed in the **Figure 1B** and **Figure 1C**. This fact has great relevance from the pharmacological point of view on prevention potential hematotoxicity effects-mediated fibrinolysis which are considered as one of the most severe adverse reactions induced by penicillin β-lactam antibiotics, immune-mediated hemolytic anemia [5], and intravascular hemolysis induced by off-target ligand interactions with tunnel sub-cavities of fibrinogen binding sites. Because, it is well-known that the critical functions of the tunnels include: i) substrate and solute exchange, ions and water between the sub-cavities, ii) transport of intermediates of biochemical reactions between two different binding sites, and iii) maintenance of the conformational and geometric complementarity of binding sites for binding physiological substrates [41]. Afterward, in order to assess the biochemical relevance of the three main regions of the fibrinogen structure (E-region, D-region 1 and D-region 2) we examine three relevant pocket descriptors which are directly associated to the protein ability to bind potential drugs namely: i) pocket stability (Ps_fib), ii) pocket continuity (Pc_fib), and iii) pocket correlation ($Pc_{i\text{-}j}$_fib)[42]. The stability parameter of the fibrinogen binding pockets belonging to the cited regions can be calculated according to the following **equations 4** and **5**:

$$P_{s\_fib} = \{S_1, S_2, \dots, S_i, \dots, S_m\} \quad (4)$$

$$S_i = {}^{n}/_{N} \quad (5)$$

where P_{s_fib} represents the fibrinogen pocket stability; S_i corresponds to the stability of the grid formed by the pocket; the parameter m represents the number of grids in the

fibrinogen pocket, n is the number of occurrences of the i-th grid in the pocket conformations, and N is the total number of the pocket conformations. The results on the stability parameter of the fibrinogen molecule are depicted in the **Figure 2.**

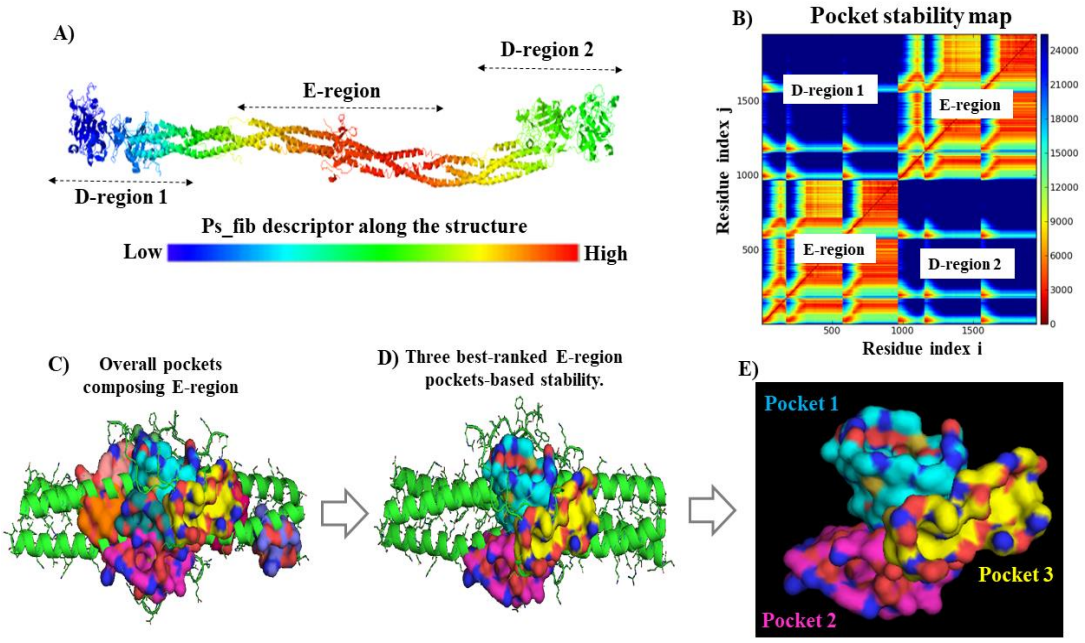


Figure 2. A) Representation of the stability properties of fibrinogen binding pockets based on the Ps_fib crystallographic molecular descriptor; from low to high stability as depicted by the bar color (from blue to orange). **B)** 2D-matrix of pocket stability map based on the Ps_fib parameter by relevant fibrinogen regions. **C)** Representation of overall pockets (10 pockets in total) composing the E-region like van der Waals surface. **D)** and **E**) van der Waals representation of the three best-ranked fibrinogen E-region pockets with the highest stability (pocket 1 > pocket 2 > pocket 3 > remaining E-region pockets >>> D-region 2 > D-region 1). Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

For the fibrinogen E-region, 10 pockets in total were detected across the E-region surface while for the D-region 1 and 2, 2 and 4 pockets were detected, respectively. See (**Figure 2A**). According to this the fibrinogen E-region suggest could has greater influence for ligand-binding affinity compared with the two D-regions. Particularly, in the E-region the central pocket 1 forms a funnel-shaped hydrophobic cavity that exhibits the greatest stability of all the E-region binding sites detected. Theoretically suggesting that the pocket 1 represents the intrinsic binding pocket for the dynamic and functional properties

of the fibrinogen protein. This pocket 1 is surrounded by the pockets 2, 3 and 4 and the stability analysis showed that the pocket 2 and 3 are slightly more stable than the pocket 4. **Figure 2C** (labelled-orange van der Waals surface). Maybe the pockets 2 and 3 provide an allosteric pathway for the antibiotic-ligand entry and exit to the main pocket 1. Besides, considering that pockets 2 and 3 cover a large surface of the central segment of E-region. The pockets (5, 7, 8) could also be considered like potential allosteric binding pockets allowing accommodate small ligand molecules. Due to their relative position respect to the funnel-shaped hydrophobic cavity that forms the main pocket 1. The remaining pockets (pockets 6,9, and 10) could act binding small molecules as several co-factors and other native substrates during the binding of thrombin into the E-region for the fibrin polymerization during the blood coagulation process based on based on its lower stability compared with the pockets 1, 2 and 3. Herein, it is important to note that the stability of the predicted pockets composing the D-regions were from medium to very low pocket stability. See **Figure 2A**.

The fibrinogen binding pocket continuity (P_{c_fib}) [42], was analyzed based on general structural attributes as appearance, merging, and volume change. The continuity descriptor is defined by the **equation 6**:

$$P_{c\_fib} = \{P_i | P_i \cap P_{ref} > 0\}, i = 1, 2, \dots, n \quad (\mathbf{6})$$

where, P_{ref} represents the evaluated pocket in the first or native conformation, P_i corresponds to the *i*-th pocket which is spatially overlapped with the P_{ref}, n represents all the predicted pockets modeled, and P_{c_fib} denotes the binding pocket continuity consecutively assembled pockets. This parameter (P_{c_fib}) is strongly depends on the proximity of the pockets, and on their flexibility properties which are associated to the different conformations that these pockets can adopt under the presence or absence of ligands (penicillin β-lactam antibiotics). Besides, the pocket continuity descriptor can be associated with the pocket correlation parameter ($P_{ci\text{-}j\_fib}$) through a correlation matrix[42]. For this instance, we reinforce our hypothesis about the greatest relevance of the fibrinogen E-region compared with the two D-regions 1 and 2. Due to the high structural inter-connection between the pockets that form the E-region (i.e., for the 10 binding pockets) compared with the less quantity and low structural communication observed for both D-regions. Besides, considering the strong correlation based on the high values for the parameter **(**$P_{ci\text{-}j\_fib}$**)** for the E-region compared with both D-regions 1

and 2. Where the pocket correlation parameter **($P_{ci\text{-}j\_fib}$)** for both fibrinogen D-regions shows several parts with moderate correlation as average values of $P_{ci\text{-}j\_fib}$ (colored-white) to full uncorrelated low $P_{ci\text{-}j\_fib}$ values (colored-blue), see **Figure 3.**

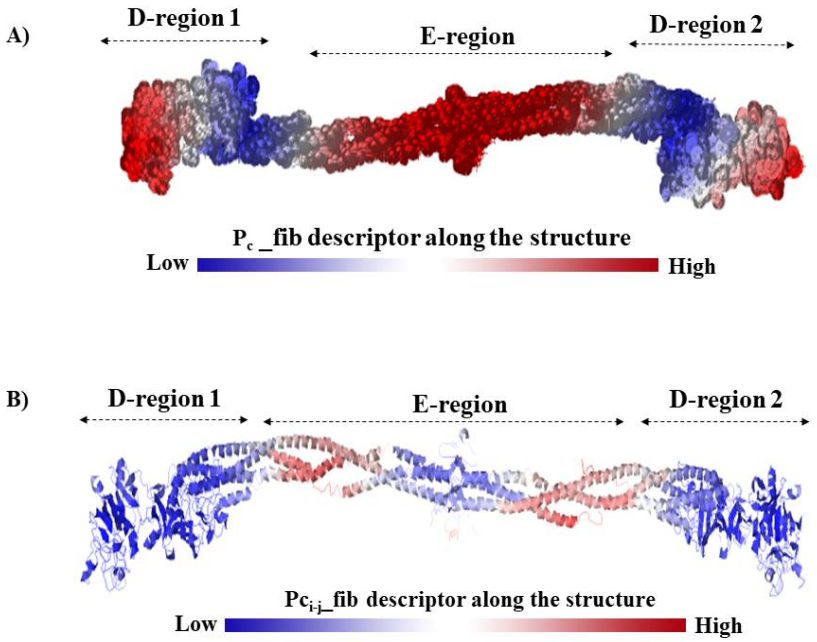


Figure 3. A) Representation of the pocket continuity crystallographic molecular descriptor (P_{c_fib}) by fibrinogen regions (E-region, D-region 1 and D-region 2). **B)** Representation of the pocket correlation crystallographic molecular descriptor ($P_{ci\text{-}j\_fib}$) for the evaluated fibrinogen regions. The bar color is to qualitatively represent the values of the crystallographic molecular descriptor evaluated as low, average, and high values according to the color intensity (blue-to-white-red transitions). Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

6.3.2 Molecular Docking Simulations

Considering all the above we performed the molecular docking simulations focus our attention on the potential docking interaction between the penicillin β-lactam antibiotics (cloxacillin and dicloxacillin) with the fibrinogen E-region, see **Figure 4**.

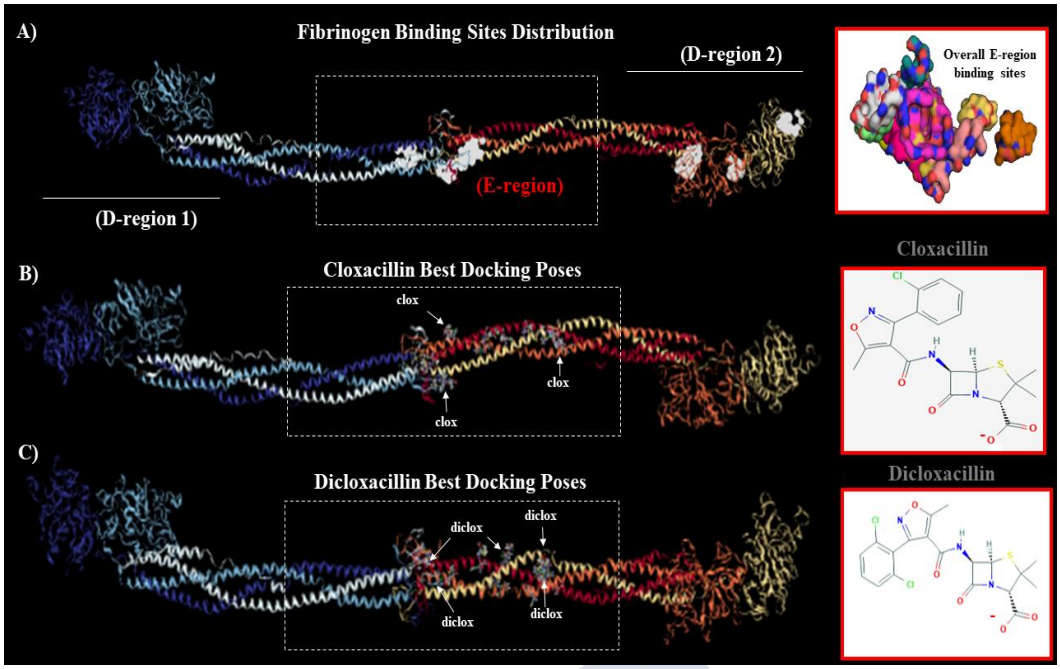


Figure 4. **A)** Fibrinogen binding sites distribution from more to less densely located by relevant regions as white volumetric surfaces (E-region > D-region 2 > D-region 1). On the far right is represented overall the binding sites that form the fibrinogen E-region. **B)** Representation of the best crystallographic docking poses obtained for the fibrinogen-cloxacillin complexes, and **C)** representation of the best crystallographic docking poses obtained for the fibrinogen-dicloxacillin complexes showing greater affinity for the region E. On the far right is represented the chemical structure of both β-lactam antibiotics cloxacillin and dicloxacillin, respectively. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

From the structural point of view, the fibrinogen E-region binding-sites are highly complex systems which contain a large variety of clefts, grooves, protrusions inner the main funnel-hydrophobic cavity (pocket 1) formed by the γ-γ and α-α crosslinking from Aα and Bβ chains composing the thrombin binding site [40]. Under the presence of ligands (cloxacillin and dicloxacillin), some critical residues can be perturbed by modifications in the symmetry architecture of the Bβ-γ/Bβ-γ dimeric domain of E-region and could induce potential hematotoxicity effects-mediated like fibrinolysis when these antibiotics are parenterally administered [5]. In general terms, the obtained results suggest that the antibiotic interactions fit with a spontaneous thermodynamic process (ΔG bind < 0 kcal/mol). Pointing the formation of stable docking complexes with the several predicted fibrinogen E-region binding sites. As showed in the **Figure 4B** and **Figure 4C** where the best crystallographic docking poses were centrally located around the funnel-

shaped hydrophobic cavity, namely the main pocket 1 previously identified like the most stable pocket. Then, aiming to deepen the modeling results, we carried out a clustered docking analysis involving the three fibrinogen regions studied, see **Figure 5**.

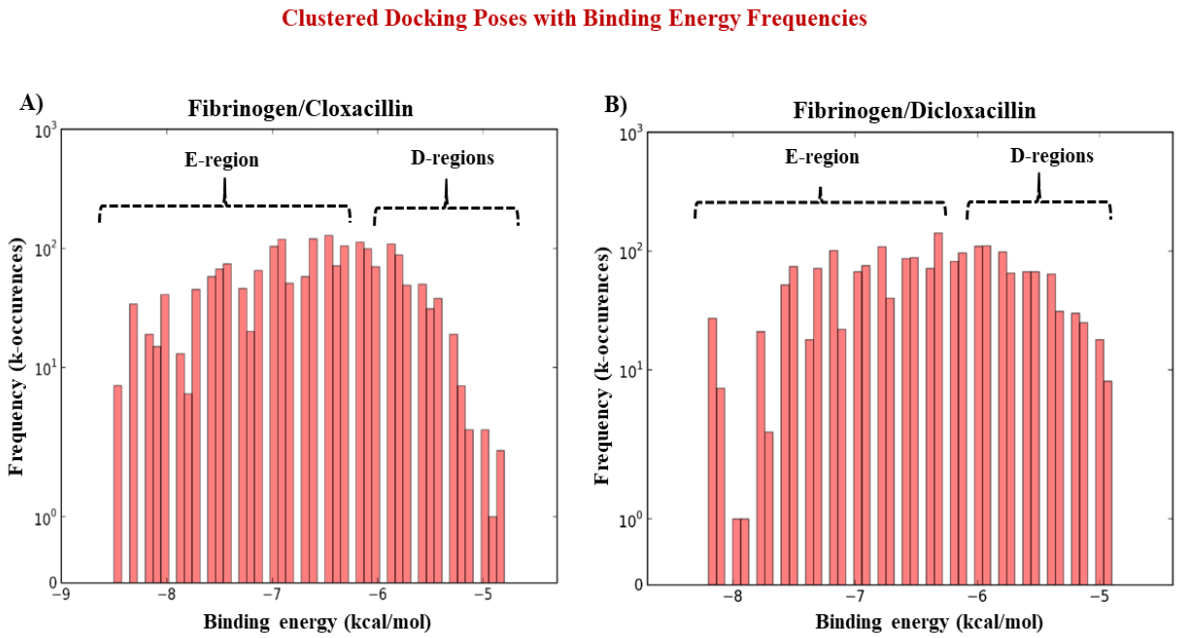


Figure 5. Frequency distribution of total binding energies from the docking complexes obtained for both β-lactam antibiotics by relevant regions (E-region and D-regions) as: **A)** fibrinogen-cloxacillin complex and **B)** fibrinogen-dicloxacillin complex. Herein, the x-axis represents the total affinities (ΔG_{bind_total}) and y-axis ordinate represents the frequencies (k-*occurrences) of* ΔG_{bind_total} values (kcal/mol) obtained for all docked poses by fibrinogen regions (*pink bars).* Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

The results showed in the **Figure 5** corroborates that the antibiotic binding interactions occur with a higher binding probability in the fibrinogen E-region compared with the two D-regions regarding the count of binding affinity frequencies (k-occurrences)[43]. In fact, the calculated values of the k-occurrences in fibrinogen E-region (k = 10-100) were higher compared with the values of k-occurrences of interactions for both penicillin β-lactam antibiotics in the D-regions (k = 20-50), approximately, which denote the preference of the cloxacillin and dicloxacillin by the central fibrinogen E-region, in both cases.

In order to theoretically explain the pharmacodynamic mechanisms of interactions for cloxacillin and dicloxacillin with the fibrinogen E-region and detect potential similarities and differences from the structure-fibrinogen activity relationships point of view a

detailed 3D/2D-lig-plot interaction diagrams were performed for both antibiotics [31]. See **Figure 6**.

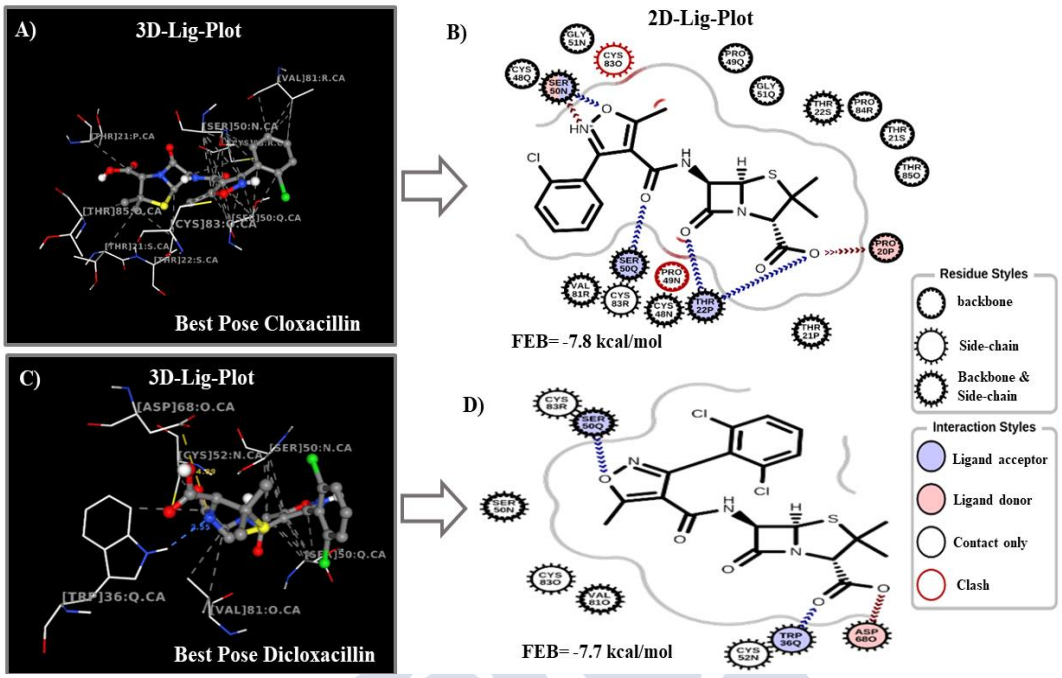


Figure 6. Representation of 3D and 2D lig-plots interaction diagrams for the best β-lactam antibiotic binding-poses obtained in the fibrinogen E-region like: **A-B**) fibrinogen-cloxacillin and **C-D**) fibrinogen-dicloxacillin. On the far right is depicted the different types of hydrophobic interactions of the ligands with the fibrinogen E-region target residues and chains as: ligand/side-chain interactions, ligand/backbone-side chain interactions, and electrostatic interactions as: ligand/backbone acceptor, ligand/backbone donor, ligand/side-chain acceptor, ligand/side-chain donor; the gray outline represents the proximity contour of the 2D lig-plots interaction diagrams. Please, note the different crystallographic orientation of the CL-atom(s) of the cloxicillin and dicloxicillin during the interactions marked in green in the obtained 3D-lig-plot diagrams. More details can be found in **Figure S1** and **Figure S2**. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

Herein, it is possible to note that, the cloxacillin molecule involves a greater number of E-region target-residues covering a greater area in the E-region biophysical environment (pocket 1) compared with the dicloxacillin molecule; despite the similar structural size of both antibiotics. This fact could be explained due to in the case of cloxacillin, the best binding pose (cloxacillin/E-region) shows an extended shape. However, the dicloxacillin ligand adopts a more closed conformation during the docking interaction. An overview

of the results of 3D/2D-lig-plot diagrams revealed that the most relevant antibiotic binding interactions with the fibrinogen E-region (pocket 1) are mainly based on hydrophobic (C···C)-backbone-side-chain non-covalent interactions, involving a greater number of these for the cloxacillin >> dicloxacillin. In this regard, we suggest that the presence of less quantity of hydrophobic (C···C)-backbone-side-chain non-covalent interactions for the dicloxacillin could promotes a decrease of the thermodynamic destabilizing effects on the residue side-chain packing of the E-region (pocket 1) [31, 44-46], potentially inducing a minor quantity of local perturbations over the interacting chains composing the E-region and without affect the flexibility properties and ligand-binding properties of the fibrinogen. Consequently, this dicloxacillin chemical modification (Cl-atom addition) contributes to prevent potential immune-mediated hemolytic anemia, intravascular hemolysis, and the most representative the effects namely fibrinolysis more efficiently way than the cloxacillin molecule. As showed in the 2D-lig-plot diagrams (panel B and D) the β-lactam ring is the most important toxicophoric fragment inducing side effects (off-target interactions) of β-lactam antibiotics (cloxacillin >> dicloxacillin) on fibrinogen. Toxicophoric groups, correspond to chemical motifs within a molecule responsible for the appearance of fibrinolytic side-effects leading defects in the blood clotting process[5].

From the therapeutic point of view, it is important highlight that the β-lactam ring is the pharmacophoric fragment responsible for the specific antibiotic action for both evaluated drugs (*i.e.*, in the absence of off-target interactions with fibrinogen blood plasma protein)[5]. In this context, the 2D-lig-plot diagrams show that the β-lactam ring belonging to the cloxacillin exhibits a large number of hydrophobic (C···C)-backbone-side-chain non-covalent contacts with fibrinogen E-region residues and chains (PRO49:Q, GLY51:Q, THR22:P:S, PRO84:R, THR21:P:S, THR85:O:R) compared with the dicloxacillin where the β-lactam ring has a unique hydrophobic (C···C)-backbone-side-chain non-covalent contact with the fibrinogen E-region residues and chains (CYS52:N). According to this, we could hypothesize several events associated with the β-lactam ring linked to the loss of selectivity/specificity, and greater potential of the cloxacillin molecule respect to dicloxacillin to induce potential hematotoxicity (fibrinolytic side-effects) like: i) total or partial perturbation of the cys-stereochemical orientation of the H-atom pair of the β-lactam ring, ii) total or partial loss of the strain in the bicyclic system that form the four membered β-lactam ring. Particularly, the β-lactam carbonyl group must be placed in the correct orientation to provide the strain required for

the attack-based acylation reaction of the β-lactam carbonyl group with the target-enzyme family (transpeptidases) and to remain the β-lactam ring stability against the action of β-lactamases. Besides, an efficient rational drug-design of penicillin β-lactam antibiotics with optimal benefit/risk ratio requires of the amidic character of the β-lactam carbonyl group to be reduced, to enable the acylation of the active nucleophilic serine residue (SER) of the transpeptidases. Then, slight distortions in the β-lactam carbonyl group orientation induced by the interaction with the fibrinogen protein (i.e., regulatory serine residue from fibrinogen E-region) could decrease the β-lactam ring strain opening the β-lactam ring and abolish its pharmacological properties (antibiotic effects). In this concern, it is well-known that the higher strain in the β-lactam ring increase the antibiotic activity and decrease the fibrinolytic hematotoxicity [5].

A similar analysis based on the orientation of chemical pharmacophoric groups can be performed for the CH3-groups composing the thiazolidine-5-membered nitrogen saturated ring which could be partially perturbed in their pharmacological trans-cys-stereochemical orientation after hydrophobic (C⋯C)-backbone-side-chain non-covalent contacts with E-region residues and chains (PRO84:R, THR21:S, THR85:O), this perturbation conformational event not occur for the dicloxacillin molecule. Besides, 2D-lig-plot diagrams show a greater presence of backbone-side-chain acceptor interactions > backbone-side-chain donor interactions happening in the oxygen atoms that form part of the β-lactam ring of both antibiotics (cloxacillin > dicloxacillin). In the case of cloxacillin the backbone-side-chain donor interaction occurs between the ionized oxygen atom with the PRO20:P and with the ASP68:O for the dicloxacillin. This oxygen atom forms part of the carboxylic group of cloxacillin and is important to selectively bind to positively charged nitrogens of the lysine from bacteria transpeptidase enzyme. By the other hand, we show that the carbonyl group composing the thiazolidine-5-membered nitrogen saturated-ring and the amide-oxygen that form part of the neighboring acylamino side-chain of the cloxacillin strongly interact with the regulatory fibrinogen E-region residues (SER50:Q and TRH22:P) through backbone-side-chain acceptor interactions. However, in the case of dicloxacillin molecule the a single backbone-side-chain acceptor interaction occurs between the carbonyl group composing the thiazolidine-5-membered nitrogen saturated-ring with a non-regulatory fibrinogen E-region residue (TRP36:Q) which significantly decreases the occurrence of adverse effects (hematotoxicity-based fibrinolysis and coagulation defects) for the dicloxacillin. It is important to note that electrophilic carbonyl groups in both ligands favor nucleophilic attack to NH groups of

regulatory fibrinogen residues like SER >> THR which present a crucial role in the phosphorylation processes during the thrombin bind to E-region which has critical importance for the blood coagulation process (fibrin polymerization). See **Figure S3**.

In addition, an interesting backbone-side-chain acceptor/donor triangular-shaped interaction between the O-atom and N-atom from the 5-membered isoxazolyl ring side-chain of the cloxacillin with high affinity for the regulatory fibrinogen E-region residue (SER50: N) was observed. This backbone-side-chain acceptor/donor triangular-shaped interaction was not observed for the dicloxacillin which showed weaker interaction involving only the regulatory E-region residue (SER50: Q). In this regard, it is important to highlight that the 5-membered isoxazolyl ring side-chain acts as steric shield but also favor the electron-withdrawing giving the structure acid stability modulating the β-lactam penicillinase resistance phenomena and hematotoxicity. According to this, the semi-synthetic incorporation of an additional halogen CL-atom in the dicloxacillin molecule, respect to cloxacillin significantly reduces the occurrence of primary and secondary fibrinolysis events respect to cloxacillin molecule through the action of strong negative inductive effects on the 5-membered isoxazolyl ring. This inductive-effect induce a notable weakening of the bonds between a hetero-atom (O, N) and a H-atom. Then, the electron-withdrawing group increases the acidity of the O-H bond of an acid (weakening of the O-H bond), but decreases the basicity of a nitrogen atom in an amine by decreasing the electron density in the free doublet allowing a more efficient modulation of the interaction properties when these β-lactam antibiotics are parenterally administrated. Besides, the differences observed in the relative docking pose-orientation in the E-region binding site for the CL-atoms of the dicloxacillin respect to single CL-atom from cloxacillin apparently is very important to maintain its pharmacological efficacy without compromising binding affinity of the formed docking complexes with very close FEB values as FEB (cloxacillin/E-region) = - 8.7 kcal/mol and FEB (dicloxacillin/E-region) = -7.7 kcal/mol.

Following the modeling procedures, we determine the total binding affinity (ΔG_{bind_total}) for each ligand, as well as their thermodynamic contributions (ΔG_{Gauss1}, ΔG_{Gauss2} , $\Delta G_{repulsion}$, $\Delta G_{H\text{-}bond}$, $\Delta G_{hydrophobic}$ and $\Delta G_{rotational}$ [47, 48]). See **Figure 7**.

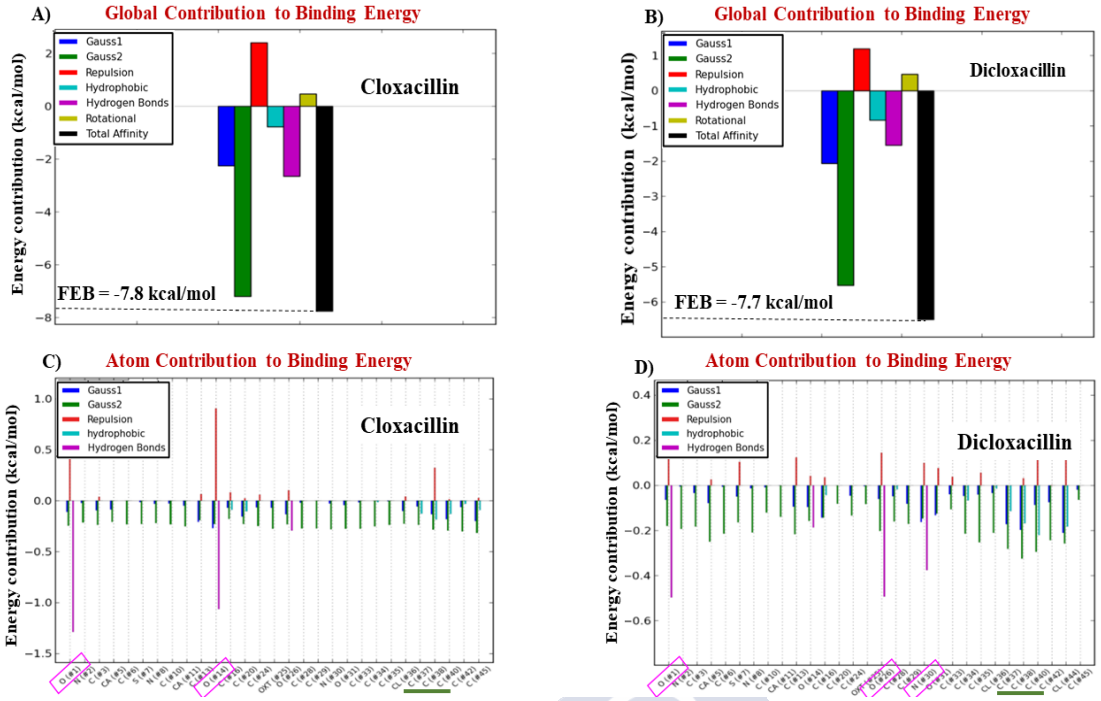


Figure 7. On the top, graphical breakdown of the global energy contribution to the total binding affinity (ΔG_{bind_total}) values (kcal/mol) obtained from the interaction of each β-lactam antibiotics with fibrinogen E-region like: **A)** fibrinogen-cloxacillin (ΔG_{bind_total} = -7.80 kcal/mol) and **B)** fibrinogen-dicloxacillin (ΔG_{bind_total} = - 7.70 kcal/mol). Herein, different color-labels of the bars were used to indicate the type of thermodynamic contribution like: attractive energy dispersion based on two gaussian functions (Gauss1 and Gauss2 like blue and green, respectively), repulsion interaction energy (red), hydrophobic interaction energy (light blue), hydrogen bond energy (purple), rotational energy (yellow) and total binding affinity (black). On the bottom, graphical representation based on per atom energy contributions for each β-lactam antibiotic tested as: **C)** fibrinogen-cloxacillin and **D)** fibrinogen-dicloxacillin. Please, note that the symbol (#) *x-axis* was used only for labeling purposes not fitting with the atom-position in the β-lactam antibiotic ligand structures. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

An overview of the molecular docking results showed in **Figure 7**, reveals that the most prevalent energy contributions to the total binding affinity ($\Delta G_{bind\_total} = \Delta G_{bind\_Clox} + \Delta G_{bind\_Diclox}$) from both β-lactam antibiotic drugs (ΔG_{bind_Clox} and ΔG_{bind_Diclox}) were associated to the contributions of the pair of steric Gaussian molecular mechanism distance-dependent terms (Gauss 2 and Gauss 1), which can be defined as attractive terms

of London dispersion forces[44-48]. In this sense, the Gauss 2 energy contribution provides relevant information on the influence of attractive binding energies of the docking complexes formed by the best docking pose of cloxacillin and dicloxacillin drugs with the fibrinogen E-region binding site (pocket 1). Herein, the procedure for obtaining the total binding affinity is according to the following **equations 7** and **8**:

$$\Delta G_{bind\_Clox} = (\Delta G_{Gauss1_{Clox}} + \Delta G_{Gauss2\_Clox)} + \Delta G_{repulsion_{Clox}} + \Delta G_{H-bond_{Clox}} + \\ + \Delta G_{hydrophobic\_Clox} + \Delta G_{rot\_Clox} \quad (7)$$

$$\Delta G_{bind\_Diclox} = (\Delta G_{Gauss1_{Diclox}} + \Delta G_{Gauss2\_Diclox)} + \Delta G_{repulsion_{Diclox}} + \Delta G_{H-bond_{Diclox}} + \\ + \Delta G_{hydrophobic\_Diclox} + \Delta G_{rot\_Diclox} \quad (8)$$

The results obtained for the Gauss 2 forces show the influence of strongest interactions with the fibrinogen E-region (pocket 1) based on non-covalent hydrophobic and non-electrostatic attractive energy contributions for both ligands evaluated and fitting well with the previous results. Herein, we strongly suggest that the high influence exerted by the Gauss 2 energy force ((Gauss 2_Clox > Gauss 2_Diclox), is due to temporary dipoles formation (fluctuating dipole–induced dipole) between the β-lactam antibiotics and the E-region binding site (pocket 1). Which present a greater number of tunnels and binding interaction surface increasing the probability-based frequency (k-occurrence) in the E-region compared with the two D-regions. However, the energetic contributions based on ΔG repulsive and ΔG rotational forces revealed a similar pattern of interaction with the E-region binding site for both ligands evaluated and were considered with low relevance for the stabilization of the obtained docking complexes from the quantitative point of view in accordance to a non-spontaneous thermodynamic process (ΔG repulsion and ΔG rot > 0 kcal/mol).

Following, we carried out an additional docking simulation in order to identify the individual atom contribution to the total binding affinity (ΔG bind) of each β-lactam antibiotic (ΔG_{bind_Clox} and ΔG_{bind_Diclox}). To this end, a per atom energy contribution analysis was performed regarding the cited thermodynamic energies (ΔG_{Gauss1}, ΔG_{Gauss2}, $\Delta G_{repulsion}$, $\Delta G_{H\text{-}bond}$, $\Delta G_{hydrophobic}$ and $\Delta G_{rotational}$) (**Figures 7; C, D)**. This approach allows a detailed exploration on the contribution of the individual atoms of the β-lactam antibiotic ligands interacting with the fibrinogen E-region binding site (pocket 1) [40-44].

The results obtained pointing that the docking complex stabilization was mainly based on $\Delta G_{H\text{-}bond}$ interaction atom energy contribution for both ligands evaluated ($\Delta G_{H\text{-}bond\_Clox} > \Delta G_{H\text{-}bond\_Diclox}$). Specifically, we observe that, the maximum atom contribution in the case of cloxacillin was provided by the H-bond interaction, generated by the cloxacillin O-atom-(#1) which belong to the four membered β-lactam ring with an H--O group from the electron-acceptor regulatory residue THR22:P:CA having a maximum atom-energy contribution around -1.4 kcal/mol. While for case of dicloxacillin the H-bond atom contribution occurs in a different position involving the ionized oxygen atom (#1) from the carboxylic group of the thiazolidine-5-membered nitrogen saturated ring with a non-regulatory residue PRO20:P with a lower $\Delta G_{H\text{-}bond}$ atom-energy around -0.5 kcal/mol. Another important difference was identified for the second best-ranked $\Delta G_{H\text{-}bond\_clox}$ atom energy contribution of cloxacillin which involves a H-bond interaction of the oxygen atom (#14) of the β-lactam ring with the regulatory residue THR22:P:CA with energy value ≈ -1 kcal/mol. However, in the case of the oxygen atom (#14) of the dicloxacillin this $\Delta G_{H\text{-}bond\_Diclox}$ atom-energy was negligible with value around -0.2 kcal/mol. Interestingly, we observed an additional $\Delta G_{H\text{-}bond\_Diclox}$ atom-energy contribution involving the oxygen atom (#26) and the nitrogen atom (#30) of the dicloxacillin around -0.45 to -0.35 kcal/mol which are totally absent when we consider the same atom positions (oxygen atom (#26) and the nitrogen atom (#30)) for the cloxacillin molecule. Furthermore, we note the presence of a cluster of hydrophobic C-atoms (#37, #38, #40 and #42) from the dicloxacillin molecules exhibit a higher atom-energy contribution based on the steric Gaussian molecular distance-dependent (Gauss 2_Diclox ≈ -0.38) compared with the energy contribution for the same cloxacillin atom positions with a lower atom-energy contribution (Gauss 2_Clox value ≈ - 0.22 kcal/mol)[43-49]. We suggest that this energetic difference could favors a more efficient pharmacological modulation based on Gauss 2 atom energy contribution (dicloxacillin > cloxacillin) toward to reduce the potential hematotoxicty effects (fibrinolysis) and avoid antibiotic resistance mechanisms.

6.3.3 LPRS maps based on the most relevant interacting E-region chains

By the other hand, we could suggest that the molecular docking mechanism of the evaluated β-lactam antibiotics with the fibrinogen molecule (E-region) is likely-based on local-perturbations (interactions) with regulatory chains (N and Q) of E-region. For this instance, the chains N and Q could be considered like critical targets to establish docking

complexes with high binding affinity with the evaluated β-lactam antibiotics because: i) N and Q chains contain the druggable-regulatory residues (SER50:N and SER50:Q) with high propension to interact with β-lactam antibiotics (refer to **Figure 6**, **B** and **D**), ii) the relative positions for these regulatory residues SER50:N and SER50:Q allow functional conformational changes of E-region. Because, are critically located close to loop-linked to helices (H2 of the Aα and Bβ chains) in the opening of the funnel-shaped hydrophobic cavity (main pocket 1) facing each other in the thrombin binding-domain (**Figure S3**) and iii) these regulatory residues SER50:N and SER50:Q showed relevant backbone-side-chain acceptor/donor interactions with the pharmacophoric groups (5-membered isoxazolyl ring side-chain) responsible for the biochemical modulation of the β-lactam penicillinase resistance and prevent potential hematotoxicity effects-mediated fibrinolysis for both antibiotic drugs [5]. For this instance, we build the local-perturbation response scanning maps (LPRS maps) for the antibiotic target N and Q chains to model local perturbations induced by the β-lactam antibiotics. It is well-known that the transitions between the E-region functional states (open and close) are dependent on its flexibility properties and require the biophysical and biochemical regulation favored by the network of allosteric communication (inter-residue network communication). LPRS maps approaches are defined from elastic network models (ENM) which offer a computationally attractive way to reduce the dimensionality of the docking problem: do not require time-consuming simulations, benefitting from the high speed and practicality[50-57]. The LPRS maps based ENM models have been widely reported to study relevant conformational changes of several proteins (as fibrinogen) at the atomistic and molecular level. Because, it is well-known that the first elastic normal modes could describe a large number the conformational differences between the protein unbound state (as unoccupied E-region) and the bound state (β-lactam antibiotics interacting with E-region). In this regard, one of the most important steps is the determination of the degree of collectivity of the system evaluated [50-57]. The degree of collectivity (K_k) of a given mode (k) determines the collective protein motions like an extent to which the structural elements (block of consecutive residues) move together in a given conformational mode [53-55]. Then, a high degree of collectivity ($K_k \approx 1$) represents a highly cooperative conformational changes strictly dependent on the low-frequency normal modes (k) that exhibit larger amplitudes of correlated motions for a large number of C-(α) atoms affecting partially or the whole structure describing global rotations-translations of consecutive residue blocks by conformational anisotropic movements. While, high-

frequency modes exhibit small amplitude and localized motions (as bending and bond stretching) which are usually associated with low collectivity ($K_k \approx 0$) from residue blocks in the polypeptide chain. In the present work, the anisotropic network models-based on rotations-translations blocks approach (ANM-RTB model) was used because it efficiently provides a good biophysical approach to describe collective motions of atoms as rigid-body translations/rotations of blocks (RTB) [25, 53-57]. Herein, the protein (E-region chains N and Q) was splitted into n-blocks, being a block composed by one or several consecutive residues clustered like rotation-translation-blocks to assess the impact under the absence or presence of docking interactions with the β-lactam antibiotics (cloxacillin and dicloxacillin). To this end, the ANM-RTB models of the interacting N and Q chains were obtained at the low-frequency normal modes as a linear combination of the rotations and translations movements of residue blocks [25, 38, 53-57]. Specifically, we use the C-(α) atoms of the most relevant interacting E-region chains (N and Q) setting a distance cutoff equal to 10 Å. Then, we perform the energy minimization of the input chains N and Q following the ANM standard procedures. For this instance, all the C-(α) atoms composing the cited chain structures (including the regulatory target residues SER50:N and SER50:Q) were energetically minimized and used to compute the ENM-Hessian matrix. ENM-Hessian matrix allows to evaluate multidimensional potential energy surfaces for biophysical systems (like E-region polypeptide chains N and Q) of coupled C-(α) atoms of residues, where the matrix elements are the second derivatives of the harmonic potential (V). For this instance, for the unbound state for the evaluated receptor chains (R: N and/or Q chains) the Hessian matrix can be defined according to the following general **equation 9**:

$$V_R = \begin{bmatrix} \left(\frac{\partial^2 V_R}{\partial X_i \, \partial X_j}\right) & \left(\frac{\partial^2 V_R}{\partial Y_i \, \partial X_j}\right) & \left(\frac{\partial^2 V_R}{\partial Z_i \, \partial X_j}\right) \\ \left(\frac{\partial^2 V_R}{\partial X_i \, \partial Y_j}\right) & \left(\frac{\partial^2 V_R}{\partial Y_i \, \partial Y_j}\right) & \left(\frac{\partial^2 V_R}{\partial Z_i \, \partial Y_j}\right) \\ \left(\frac{\partial^2 V_R}{\partial X_i \, \partial Z_j}\right) & \left(\frac{\partial^2 V_R}{\partial Y_i \, \partial Z_j}\right) & \left(\frac{\partial^2 V_R}{\partial Z_i \, \partial Z_j}\right) \end{bmatrix} \quad (\mathbf{9})$$

Relevant perturbations which affect the E-region collective motions for the bound state can be also modeled as rigid-body translations/rotations of blocks (RTB) of C-(α) atoms by considering the allosteric signal propagations (local perturbations) released from the critical regulatory-sensor residues (i-residues: SER50:N and SER50:Q) to other consecutive neighboring residues (j-effectors) in the evaluated chain. Herein, the receptor

chains (R: N and Q chains) are represented as a projection matrix (P) of an atomistic C-(α) $3N_a$ x $3N_a$ Hessian matrix (H_b) into a small $6n_B \times 6n_B$ block matrix being given by $H_b = \dot{V}_{(R)_{ij}}^{b} = P^t H P$ [25, 53-57]. Where, P represents an orthogonal $3N_a$ x $6n_B$ matrix composing by the vectors associated with the local rotation and translation motions-based perturbations of each block (B), N_a represents the number of C-(α) atoms in the evaluated chains (N or Q) with the corresponding number of blocks (n_B) chosen for the evaluated chain in question. Here, the query normal low-frequency modes (k) are the eigenvectors of the cited above Hessian matrix which holds the anisotropic information's on the collective fluctuations of inter-residue network communication observed in crystallographic fibrinogen E-region structure for the N- or C-termini. The $6n_B$ vectors belonging to the projection matrix (P) can be express in a new form as $\dot{V}_{(R-L)_{ij}}^{b} = P_i^t(\dot{V}_{(R-L)_{ij}}^{b})P_j$ for each diagonal block for the bound state. Next, we can rewrite the **equation 9** to study the changes in the conformational dynamic of our systems (chains N and Q) upon docking interaction (bound state) with the ligands (L: penicillin β-lactam antibiotics cloxacillin and dicloxacillin) in the local 3D-space according to the **equation 10**:

$$\dot{V}_{R-L} = \begin{bmatrix} \left(\frac{\partial^2 V_R}{\partial X_i \partial X_j} + \frac{\partial^2 V_L}{\partial X_i \partial X_j}\right) & \left(\frac{\partial^2 V_R}{\partial Y_i \partial X_j} + \frac{\partial^2 V_L}{\partial Y_i \partial X_j}\right) & \left(\frac{\partial^2 V_R}{\partial Z_i \partial X_j} + \frac{\partial^2 V_L}{\partial Z_i \partial X_j}\right) \\ \left(\frac{\partial^2 V_R}{\partial X_i \partial Y_j} + \frac{\partial^2 V_L}{\partial X_i \partial Y_j}\right) & \left(\frac{\partial^2 V_R}{\partial Y_i \partial Y_j} + \frac{\partial^2 V_L}{\partial Y_i \partial Y_j}\right) & \left(\frac{\partial^2 V_R}{\partial Z_i \partial Y_j} + \frac{\partial^2 V_L}{\partial Z_i \partial Y_j}\right) \\ \left(\frac{\partial^2 V_R}{\partial X_i \partial Z_j} + \frac{\partial^2 V_L}{\partial X_i \partial Z_j}\right) & \left(\frac{\partial^2 V_R}{\partial Y_i \partial Z_j} + \frac{\partial^2 V_L}{\partial Y_i \partial Z_j}\right) & \left(\frac{\partial^2 V_R}{\partial Z_i \partial Z_j} + \frac{\partial^2 V_L}{\partial Z_i \partial Z_j}\right) \end{bmatrix} \quad (\mathbf{10})$$

In this context, for calculating the normal modes we use Dynomics software (DIAGQ routine) [25, 53-57]. Which automatically allows the diagonalization of the obtained ENM-Hessian matrix for the minimized unbound chain structures (by diagonalizing the N and/or Q Hessian matrix) and in the presence of β-lactam antibiotics by diagonalizing the receptor–ligand docking complexes (R-L Hessian matrices) namely: chain N+cloxacillin, chain N+dicloxacillin, chain Q+cloxacillin, chain Q+dicloxacillin, chain N-Q+cloxacillin, chain N-Q+dicloxacillin, considering in all the cases, the best crystallographic docking poses. Then, we tackle the construction of the local perturbation response scanning maps (LPRS maps) based on the low-frequency modes (k) [25, 53-57]. The LPRS maps generates a comprehensive visualizations which allows evaluate

allosteric signal propagations in response to external perturbations (i.e., as induced by the β-lactam antibiotic interactions from the regulatory residues SER50:N and SER50:Q). Following this idea, we suggest that conformational changes in the intrinsic chain dynamics, as well as, perturbations in the flexibility properties of the N and Q chains could be associated to potential side effects of cloxacillin > dicloxacillin as anomalous binding of the thrombin molecule in the E-region due to loss of allosteric cooperative mechanisms and conformational coupling between normal modes (degree of collectivity) for these target chains under β-lactam antibiotic docking interactions. The degree of collectivity (K_k) for a given block of consecutive residues is proportional to the exponential of the ‘‘information entropy’’ of the eigenvector (k-mode) and can be defined according to Bruschwiler's **equation 11**:

$$\text{Collectivity\_K}_k = \frac{1}{N}\exp\left(-\sum_i^N \alpha\left(\Delta d_{ij\_k}\right)^2 \log\alpha\left(\Delta d_{ij\_k}\right)^2\right) \qquad (\mathbf{11})$$

where, are included the sum of all N atoms of the molecule (fibrinogen E-region chains), the term d_{ij_k} represents the amplitude of the displacement of the atom i in the elastic normal mode k, and α is a normalization parameter which correspond to $\sum_i^N \alpha\left(\Delta d_{ij\_k}\right)^2 = 1$. Herein, $\Delta d_{ij\_k} = d_{ij\_k} - d^0_{ij\_k}$ which represents the displacement for different amplitudes between the unperturbed and the perturbed state of the displacements (distance-based local perturbations) from equilibrium represented by ($d^0_{ij\_k}$). According to this, we could represent the degree of collectivity for locally perturbed fibrinogen chains (R, receptor as N and Q chains) interacting with the β-lactam antibiotics (L, ligands as cloxacillin or dicloxacillin) by replacing the terms in the Bruschwiler's **equation 12**.

$$\text{Collectivity\_K}_{k(R-L)} = \frac{1}{N}\exp\left(-\sum_{i=1}^{N=7} \alpha(d_{R-L} - d^0_R)^2 \log\alpha\,(d_{R-L} - d^0_R)^2\right) \qquad (\mathbf{12})$$

herein, the low-frequency normal modes allow to efficiently characterize the collective motions of the relevant receptor chains N and Q which are directly associated to functional conformational changes of E-region (main pocket 1). To address the evaluation of the corresponding (3N)-XYZ-atomic displacements we use the $d_{ij}(d^0_{ij})$ which represent the distance between the conformational dynamical in equilibrium of the C-(α) atoms network using a d_{cutoff} = 7 Å which allows map the conformational response triggered from the i-sensor regulatory residues (SER50:N and SER50:Q) to j-effector

residues. In this context, the displacement vector in the (3N)-XYZ-Cartesian representation can be extracted from the aforementioned equations to represent distance-based perturbations for the evaluated conformational states (unbound and bound) for the interacting N and Q chains as:

$$\Delta d_{ij\_k} = (X_1 - X_1^0, Y_1 - Y_1^0, Z_1 - Z_1^0, \ldots X_N - X_N^0, Y_N - Y_N^0, Z_N - Z_N^0)^t \quad (\mathbf{13})$$

Next, the **equation 9** can be rewritten as:

$$V_R = \sum_{(i<j)\in N(R)} \frac{1}{2} \gamma_R \left(\Delta d_{ij\_k}\right)^t \left(\Delta d_{ij\_k}\right) \quad (\mathbf{14})$$

By the other hand, in the present study we rigorously set the degree of collectivity for the first seven low-frequency normal modes (k, from 1 to 7, see **Figure S4**). It is important to highlight that the linear combinations of normal mode eigenvectors are defined by their correlation to the van der Waals forces derived from each C-(α) atom. Which as mentioned above, are heterogeneously distributed across the overall E-region structure in the inter-residue network that form the cluster of hydrophobic sub-cavities in the fibrinogen E-region. To explore conformational changes of regulatory chains (N and Q) under the presence of β-lactam antibiotics inducing local perturbations we perform the LPRS maps in collective low-frequency modes (K_k, from 1 to 7). See **Figures 8** and **9**.

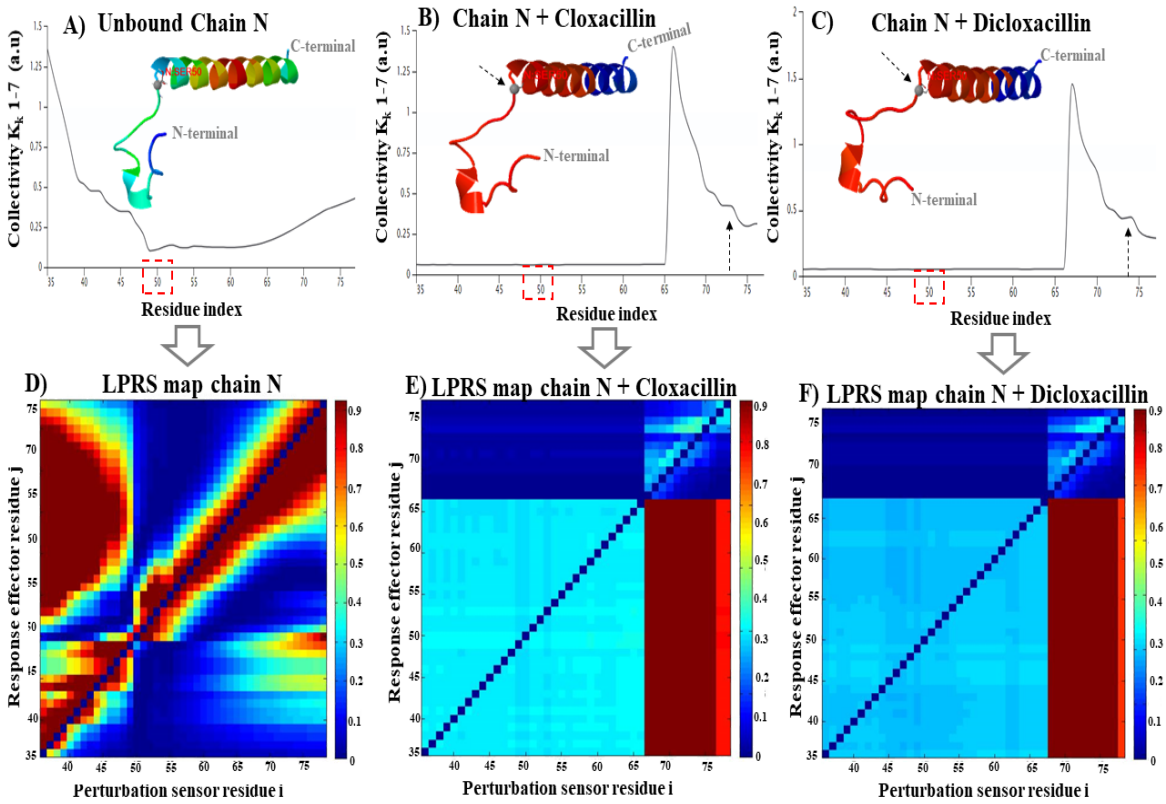


Figure 8. On the top, graphical representation of the behavior of the parameter degree of collectivity K_k (a.u) for the first seven low frequency normal modes 1-7 vs. chain N residue index, highlighting the local perturbations from the regulatory residue SER50:N (red dashed-line rectangle) under different simulation conditions as **A)** unbound chain N, **B)** chain N + cloxacillin, and **C)** chain N + dicloxacillin. The dotted black arrows represent long distance-based perturbations from the regulatory residue SER50: N. On the bottom, graphical representations of the generated local perturbation response (LPRS maps) showing the relationship between (i-j)-pairs of residues as i-sensor residues (x-axis) to j-effector residues (y-ordinates) depicted like a 2D-matrix of allosteric network communication of inter-residue (i-j pairs as off-diagonal elements) and intra-residue (i-i pairs diagonal elements). On the far right of each LPRS map, the bar-color from blue to yellow regions correspond to weak to moderate local perturbations in the inter-residue communication for (i-j)-pairs of residues, and the orange to dark red regions represent strong inter-residue local perturbations for the aforementioned simulation conditions as **D)** unbound chain N, **E)** chain N + cloxacillin, and **F)** chain N + dicloxacillin. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

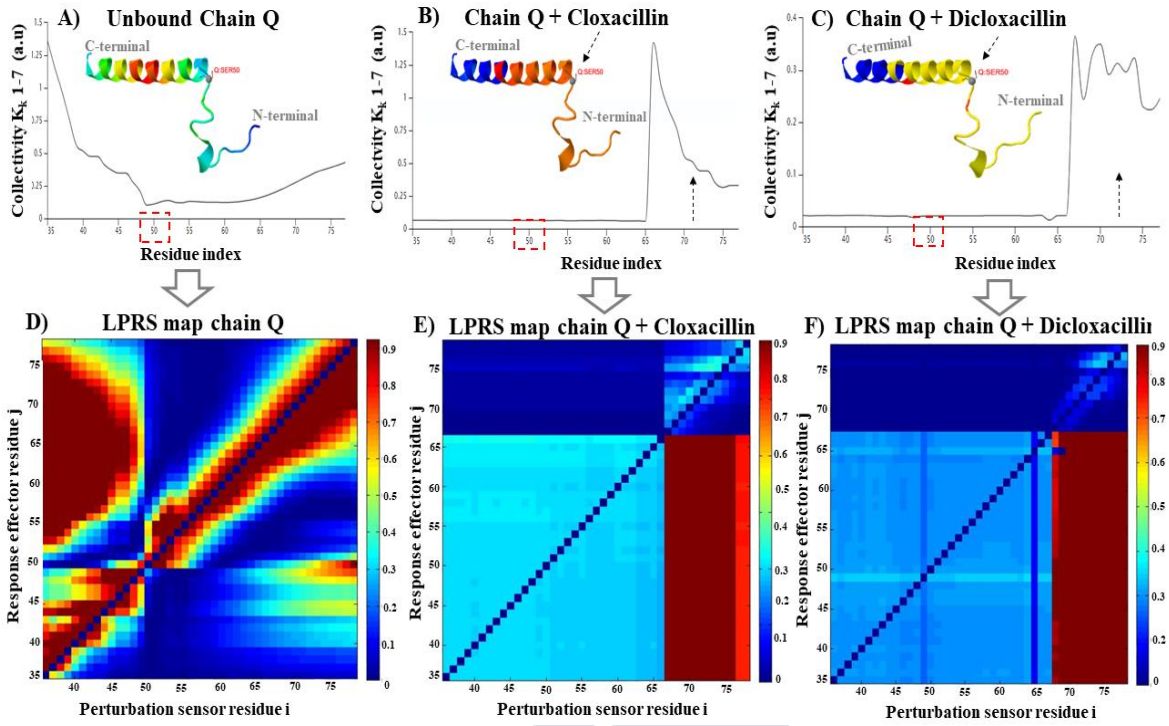


Figure 9. On the top, graphical representation of the behavior of the parameter degree of collectivity K_k (a.u) for the first seven low frequency normal modes 1-7 vs. chain Q residue index, highlighting the local perturbations from the regulatory residue SER50:Q (red dashed-line rectangle) under different simulation conditions as **A)** unbound chain Q, **B)** chain Q +cloxacillin, and **C)** chain Q + dicloxacillin. The dotted black arrows represent long distance-based perturbations from the regulatory residue SER50: Q. On the bottom, graphical representations of the generated local perturbation response (LPRS maps) showing the relationship between (i-j)-pairs of residues as i-sensor residues (x-axis) to j-effector residues (y-ordinates) depicted like a 2D-matrix of allosteric network communication of inter-residue (i-j pairs as off-diagonal elements) and intra-residue (i-i pairs diagonal elements). On the far right of each LPRS map, the bar-color from blue to yellow regions correspond to weak to moderate local perturbations in the inter-residue communication for (i-j)-pairs of residues, and the orange to dark red regions represent strong inter-residue local perturbations for the aforementioned simulation conditions as **D)** unbound chain Q, **E)** chain Q + cloxacillin, and **F)** chain Q + dicloxacillin. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

The results on LPRS maps allowed us to investigate the collectivity parameters in low-frequency modes (k, from 1 to 7), and explain the observed larger-scale movements associated to long distance-based allosteric signal propagation affecting the inter-residue network communication. In this concern, the degree of collectivity and LPRS maps pointing clearly that both antibiotics cloxacillin and dicloxacillin can significantly affect the intrinsic dynamics of the relevant regulatory chains (N and Q) that form part of the fibrinogen E-region (main pocket 1) by reducing the K_k-values when we compare the unbound state (as reference control of simulation) with the bound state, for both chains. See **Figure S5.** Although, the evaluated antibiotics belong to the same pharmacological group exhibit a different docking interaction mechanism without significantly compromising the binding affinity values (- 0.1 kcal/mol of variation between both). In this sense, we theoretically show that the cloxacillin and dicloxacillin could significantly affect (i.e., induce conformational perturbations) in a large blocks of neighboring residues from both polypeptide chains when the regulatory i-sensor residues SER50:N and SER50:Q are perturbed by the β-lactam antibiotics interaction (i.e., signal propagation generated from SER50:N and SER50:Q in both, N-terminal and C-terminal direction from the regulatory chains N and Q). Specifically, around 31 consecutive amino acid residues were conformationally perturbed with an abrupt decrease in the degree of collectivity in the C-terminal direction (See **Figure 8B**,**C** and **Figure 9B**,**C**) from GLY35:N to regulatory non-target THR65:N and from GLY35:Q to regulatory non-target THR65:Q. Consistent with phenomena of conformational rigidification in large cluster of residues according to the results of the corresponding LPRS maps (regions colored from light blue to dark blue) as showed in the **Figure 8E**,**F** and **Figure 9E**,**F** for the chains N and Q, respectively. The conformational rigidification in both chains (N and Q) in the bounded state is remarkable when we compare with the reference control represented by the unbound state for the chains N and Q as unperturbed crystallographic native conformation (including the unbound residues SER50:N and SER50:Q). See **Figure 8A**, **D** and **Figure 9 A**, **D**.

As mentioned above, a significant decrease in the degree of collectivity of the chains for K_k-values ≈ 0 is frequently associated to localized vibrational motions of a few atoms. Therefore, we strongly suggest that the β-lactam antibiotic docking interaction could induce several local perturbations based on bending and bond stretching in the inter-residue network for the target chains N and Q and theoretically promoting a significant

rigidification associated to perturbation in the flexibility properties of this chains following the order (chain N + dicloxacillin > chain N + cloxacillin >> chain N) and (chain Q + dicloxacillin > chain Q + cloxacillin >> chain Q). Apparently, a transient decrease of the degree of collectivity in the first block of residues composing the unbound state of the chains N and Q (ranging from the GLY35 to regulatory non-target THR65 residue index positions in both chains, N and Q); favor the correct accommodation and required adjustment of the native substrates (thrombin) in the E-region (pocket 1) during the coagulation process. However, from the pharmacological point of view, the potential docking interactions as promoted by parenterally administrated β-lactam antibiotic with the fibrinogen blood plasma protein, could theoretically affect the binding of thrombin associated with a significant decrease of the degree of collectivity with K_k-values ≈ 0, in these block of residues; leading appreciable changes in the conformational architecture of the E-region inter-residue network communication in the bound state generated from distance-based local perturbations initiated in SER50:N and SER50:Q. Leading potential hematotoxicity (fibrinolysis), which as mentioned above, is considered as one of the most severe adverse reactions induced by penicillin β-lactam antibiotics when are parenterally administrated [5]. In addition, we observed that both antibiotics cloxacillin and dicloxacillin promote a significant increase in the degree of collectivity with a peak, and subsequently an abrupt decrease of the collectivity parameter (K_k-values) in the residue range from Phe 65:N to SER 75:N, with subtle conformational differences as can be observed in the end of the profiles of collectivity mode vs. residue index. Where, dotted black arrows representing long distance-based perturbations (**Figure 8**, **B** and **C**) and in the corresponding LPRS maps (**Figure 8**, **E** and **F**). Where, a strong local perturbations were observed for the docking complexes formed by chain N + cloxacillin ≈ chain N + dicloxacillin involving the signal transduction of the perturbations from the block of i-sensor residues (from regulatory non-target SER67:N to Phe79:N) which trigger strong allosteric responses in the j-effector residues (ranging from GLY35:N to regulatory non-target THR65:N). Herein, non-target regulatory residues (SER and THR) are referred to residues which do not directly interact with antibiotics evaluated. By the other hand, when both antibiotics interact with the chain Q (**Figure 9**, **B** and **C**). Just the cloxacillin keeps the same pattern of K_k-value perturbation previously observed for the interaction with the N chain. However, when the regulatory target SER50:Q is perturbed by the dicloxacillin ligand the chain Q exhibit a different pattern of signal propagation across its inter-residue network communication involving the appearance of several peaks or fluctuations of the

K_k-values compared with the simulation condition for the chain Q for the complex formed by the Q-chain+cloxacillin, where the aforementioned fluctuation as peaks of the K_k-values were not observed in the residue range from PHE65:N to SER 75:N (**Figure 9**, **B** and **C**). More details can be found in **Figure S5**. Furthermore, the obtained LPRS maps for the docking complexes formed by chain Q + cloxacillin and chain Q + dicloxacillin exhibit slight differences in the pattern of signal propagation of the chain Q inter-residue network under interactions with the cited antibiotics (**Figure 9**, **E** and **F**). In this regard, we note the presence of greater number of long-distance-based local perturbations following the order: chain Q + dicloxacillin > chain Q + cloxacillin >> chain Q. Particularly, involving the signal transduction-based perturbations from the block of i-sensor residues (from regulatory non-target SER67:Q to PHE79:Q) triggering allosteric responses in j-effector residues (ranging from GLY35:Q to PHE79:Q). This, fact could be associated to the different crystallographic docking pose orientation of the CL-atom(s) of the dicloxacillin respect to cloxacillin during the interactions with E-region as previously showed 3D-lig-plot analysis **(Figure 6**). Please, also refer to **Figure S1** and **Figure S2**. To validate the intrinsic conformational dynamic of the target N and Q chains evaluated. The Ramachandran analysis was carried up as a control test to avoid obtain false positives from the results of the LPRS maps. Also the Ramachandran plots obtained from the individual regulatory residues SER50:N and SER50:Q were considered. See **Figure S6**. In this regard, allowed torsion values for *Phi vs. Psi* dihedral angles were unequivocally identified inside the Ramachandran colored purple contour lines in both chains (*i.e.,* conformationally-favored N and Q chains including their corresponding regulatory E-region residues SER50:N and SER50:Q). Corroborating that overall results obtained from the LPRS maps were carried out in the total absence of sterically disallowed residues which are usually associated with torsion dihedral angles Psi vs. Phi values ocurring outside of the Ramachandran colored purple contour lines. These results are the paramount importance because it is well known that (C···C)-backbone-side-chain interactions (bound state) are associated to fluctuations of adjacent dihedral angles Psi and Phi through collective angular motions [25, 53-57].

Besides, we carried out additional simulations by considering the 2D-matrix cross-correlation of flexibility for both interacting chains (N and Q) in the bound state and the showing that the cloxacillin and dicloxacillin can modulate the signal allosteric propagations and flexibility properties of the inter-residue network composing the E-region. See **Figure S7**.

It is important to note that, in order to validate our ANM results we perform control simulation experiments by applying well established normal mode approaches based on several DynaMut-ANM force fields namely: i) C-(α) force field derived from Amber94 all-atom potential, ii) ANM simplified spring force constant-based pair-wise distance, iii) pfANM parameter-free ANM as alternative from the ANM force field with perturbations which fall off with the square of the inter-residue distance, iv) REACH called realistic extension algorithm-based covariance Hessian parameterized with variance-covariance matrices, and lastly sdENM which employs residue-specific spring force constants based on parameterized statistical analysis of > 1500 NMR-ANM molecular ensembles [52-58]. These control ANM-validation procedures fit well with our ANM results (i.e., getting similar K_k-values for the evaluated N and Q fibrinogen E-region chains) corroborating that the proposed ANM models are robust, and can be efficiently applied to evaluate the conformational properties of our systems under the unbound and bound states with the penicillin β-lactam antibiotics.

References

[1] N. Vankadari, J.A. Wilce. Emerging COVID-19 coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26, Emerging microbes & infections, 9 (2020) 601-604.

[2] K. Tian, M. Shao, Y. Wang, J. Guan, S. Zhou. Boosting compound-protein interaction prediction by deep learning, Methods, 110 (2016) 64-72.

[3] Y. Cai, M.J. Hossain, J.-K. Hériché, A.Z. Politi, N. Walther, B. Koch, M. Wachsmuth, B. Nijmeijer, M. Kueblbeck, M. Martinic-Kavur, Experimental and computational framework for a dynamic protein atlas of human cell division, Nature, 561 (2018) 411-415.

[4] J. Alex, J. Gary. Cosolute and Crowding Effects on a Side-By-Side Protein Dimer, Biochemistry, (2017).

[5] T. Nakano, A. Terawaki, H. Arita. Influence of β-lactam antibiotics on platelets. II. In vitro effects of some β-lactam antibiotics on the biochemical response of rat platelets, Journal of pharmacobio-dynamics, 10 (1987) 408-420.

[6] N. Malangu. Poisoning: From Specific Toxic Agents to Novel Rapid and Simplified Techniques for Analysis, BoD–Books on Demand, Place Published, 2017.

[7] P. Taboada, D. Attwood, J.M. Ruso, M. García, F. Sarmiento, V. Mosquera. Effect of electrolyte on the surface and thermodynamic properties of amphiphilic penicillins, Journal of colloid and interface science, 220 (1999) 288-292.
[8] P. Taboada, D. Attwood, J.M. Ruso, M. García, F. Sarmiento, V. Mosquera. Influence of molecular structure on the ideality of mixing in micelles formed in binary mixtures of surface-active drugs. Journal of colloid and interface science, 216 (1999) 270-275.
[9] J. Ruso, D. Attwood, M. Garcia, P. Taboada, L. Varela, V. Mosquera, A study of the interaction of the amphiphilic penicillins cloxacillin and dicloxacillin with human serum albumin in aqueous solution. Langmuir: the ACS journal of surfaces and colloids, 17 (2001) 5189-5195.
[10] P. Taboada, D. Attwood, J.M. Ruso, F. Sarmiento, V. Mosquera. Self-association of amphiphilic penicillins in aqueous electrolyte solution: a light-scattering and NMR study, Langmuir : the ACS journal of surfaces and colloids, 15 (1999) 2022-2028.
[11] N. Funasaki, S. Hada, S. Neya. Self-association of penicillins in aqueous solution as revealed by gel filtration chromatography, Chemical and pharmaceutical bulletin, 42 (1994) 779-785.
[12] P. Taboada, V. Mosquera, J.M. Ruso, F. Sarmiento, M.N. Jones. Interaction between penicillins and human serum albumin: a thermodynamic study of micellar-like clusters on a protein. Langmuir : the ACS journal of surfaces and colloids, 16 (2000) 934-938.
[13] P. Taboada, V. Mosquera, J.M. Ruso, F. Sarmiento, M.N. Jones. Interaction between penicillins and human serum albumin: a ζ-potential study. Langmuir : the ACS journal of surfaces and colloids, 16 (2000) 6795-6800.
[14] N. Hassan, J.M. Ruso, P. Somasundaran. Mechanisms of fibrinogen–acebutolol interactions: Insights from DSC, CD and LS. Colloids and Surfaces B: Biointerfaces, 82 (2011) 581-587.
[15] I. Pechik, S. Yakovlev, M.W. Mosesson, G.L. Gilliland, L. Medved. Structural basis for sequential cleavage of fibrinopeptides upon fibrin assembly. Biochemistry, 45 (2006) 3588-3597.
[16] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker. PubChem substance and compound databases. Nucleic Acids Res, 44 (2016) D1202-D1213.
[17] J. Madrazo, J.H. Brown, S. Litvinovich, R. Dominguez, S. Yakovlev, L. Medved, C. Cohen. Crystal structure of the central region of bovine fibrinogen (E5 fragment) at 1.4-Å resolution, Proceedings of the National Academy of Sciences, 98 (2001) 11967-11972.

[18] A. Bratek-Skicki, P. Żeliszewska, J.M. Ruso. Fibrinogen: a journey into biotechnology, Soft Matter, 12 (2016) 8639-8653.
[19] S. Hochreiter, G. Klambauer, M. Rarey. Machine Learning in Drug Discovery, ACS Publications, 2018.
[20] X.-Y. Meng, H.-X. Zhang, M. Mezei, M. Cui. Molecular docking: a powerful approach for structure-based drug discovery, Current computer-aided drug design, 7 (2011) 146-157.
[21] S. Renner, S. Derksen, S. Radestock, F. Mörchen. Maximum common binding modes (MCBM): consensus docking scoring using multiple ligand information and interaction fingerprints, Journal of chemical information and modeling, 48 (2008) 319-332.
[22] S. Mitternacht, I.N. Berezovsky. Coherent conformational degrees of freedom as a structural basis for allosteric communication, PLoS computational biology, 7 (2011).
[23] T. Oliwa, Y. Shen. cNMA: a framework of encounter complex-based normal mode analysis to model conformational changes in protein interactions, Bioinformatics, 31 (2015) i151-i160.
[24] U. Emekli, D. Schneidman-Duhovny, H.J. Wolfson, R. Nussinov, T. Haliloglu. HingeProt: automated prediction of hinges in protein structures, Proteins: Structure, Function, and Bioinformatics, 70 (2008) 1219-1227.
[25] J.G. Greener, M.J. Sternberg. AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis, BMC bioinformatics, 16 (2015) 335.
[26] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, P. Bourne. The protein data Bank nucleic acids research. 2000, 235-242.
[27] G. Martínez-Rosell, T. Giorgino, G. De Fabritiis. PlayMolecule ProteinPrepare: A web application for protein preparation for molecular dynamics simulations, Journal of Chemical Information and Modeling, 57 (2017) 1511-1516.
[28] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, Journal of cheminformatics, 4 (2012) 17.
[29] O. Trott, A.J. Olson. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, Journal of computational chemistry, 31 (2010) 455-461.

[30] S. Forli, R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, A.J. Olson. Computational protein–ligand docking and virtual drug screening with the AutoDock suite, Nature protocols, 11 (2016) 905-919.
[31] A. Tao, Y. Huang, Y. Shinohara, M.L. Caylor, S. Pashikanti, D. Xu. ezCADD: A rapid 2D/3D visualization-enabled web modeling environment for democratizing computer-aided drug design, Journal of chemical information and modeling, 59 (2018) 18-24.
[32] V. Le Guilloux, P. Schmidtke, P. Tuffery. Fpocket: an open source platform for ligand pocket detection, BMC bioinformatics, 10 (2009) 1-11.
[33] W.P. Feinstein, M. Brylinski. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets, Journal of cheminformatics, 7 (2015) 1-10.
[34] R.A. Laskowski, M.B. Swindells. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, ACS Publications, 2011.
[35] F. Tama, F.X. Gadea, O. Marques, Y.H. Sanejouand. Building-block approach for determining low-frequency normal modes of macromolecules, Proteins: Structure, Function, and Bioinformatics, 41 (2000) 1-7.
[36] L.-W. Yang, C.-P. Chng. Coarse-grained models reveal functional dynamics-I. elastic network models–theories, comparisons and perspectives, Bioinformatics and Biology Insights, 2 (2008) BBI. S460.
[37] B.H. Lee, S. Seo, M.H. Kim, Y. Kim, S. Jo, M.-k. Choi, H. Lee, J.B. Choi, M.K. Kim. Normal mode-guided transition pathway generation in proteins, PloS one, 12 (2017) e0185658.
[38] J.G. Greener, M.J. Sternberg. AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis, BMC bioinformatics, 16 (2015) 1-7.
[39] R.J. Worthington, C. Melander. Overcoming resistance to β-lactam antibiotics, The Journal of organic chemistry, 78 (2013) 4207-4213.
[40] J.M. Kollman, L. Pandi, M.R. Sawaya, M. Riley, R.F. Doolittle. Crystal structure of human fibrinogen, Biochemistry, 48 (2009) 3877-3886.
[41] J. Stourac, O. Vavra, P. Kokkonen, J. Filipovic, G. Pinto, J. Brezovsky, J. Damborsky, D. Bednar. Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport, Nucleic Acids Res, 47 (2019) W414-W422.

[42] Z. Chen, X. Zhang, C. Peng, J. Wang, Z. Xu, K. Chen, J. Shi, W. Zhu. D3Pockets: a method and web server for systematic analysis of protein pocket dynamics, Journal of chemical information and modeling, 59 (2019) 3353-3358.
[43] I. Sánchez-Linares, H. Pérez-Sánchez, J.M. Cecilia, J.M. García. High-throughput parallel blind virtual screening using BINDSURF, BMC bioinformatics, 13 (2012) S13.
[44] N.E. Newell, Mapping side chain interactions at protein helix termini, BMC bioinformatics, 16 (2015) 231.
[45] A. Bachmann, D. Wildemann, F. Praetorius, G. Fischer, T. Kiefhaber. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction, Proceedings of the National Academy of Sciences, 108 (2011) 3952-3957.
[46] Z.-R. Xie, M.-J. Hwang. An interaction-motif-based scoring function for protein-ligand docking, BMC bioinformatics, 11 (2010) 298.
[47] R. Quiroga, M.A. Villarreal. Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening, PLOS ONE, 11 (2016) e0155183.
[48] V. Venkatraman, D.W. Ritchie. Flexible protein docking refinement using pose-dependent normal mode analysis, Proteins: Structure, Function, and Bioinformatics, 80 (2012) 2262-2274.
[49] O. Keskin, S.R. Durell, I. Bahar, R.L. Jernigan, D.G. Covell. Relating molecular flexibility to function: a case study of tubulin, Biophysical journal, 83 (2002) 663-680.
[50] C. Chennubhotla, A. Rader, L.-W. Yang, I. Bahar. Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies, Physical Biology, 2 (2005) S173.
[51] F. Tama, Y.-H. Sanejouand. Conformational change of proteins arising from normal mode calculations, Protein engineering, 14 (2001) 1-6.
[52] R. Brüschweiler. Collective protein dynamics and nuclear spin relaxation, The Journal of chemical physics, 102 (1995) 3396-3403.
[53] K. Suhre, Y.-H. Sanejouand. ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement, Nucleic Acids Res, 32 (2004) W610-W614.
[54] P. Petrone, V.S. Pande. Can conformational change be described by only a few normal modes?, Biophysical journal, 90 (2006) 1583-1593.

[55] L. Yang, G. Song, R.L. Jernigan. Protein elastic network models and the ranges of cooperativity, Proceedings of the National Academy of Sciences, 106 (2009) 12347-12352.
[56] Y. Dehouck, A.S. Mikhailov, Effective harmonic potentials: insights into the internal cooperativity and sequence-specificity of protein dynamics, PLoS Comput Biol, 9 (2013) e1003209.
[57] M. González-Durruthy, G. Scanavachi, R. Rial, Z. Liu, M.N.D.S. Cordeiro, R. Itri, J.M. Ruso, Structural and energetic evolution of fibrinogen toward to the betablocker interactions, International Journal of Biological Macromolecules, 137 (2019) 405-419.

Supplementary materials

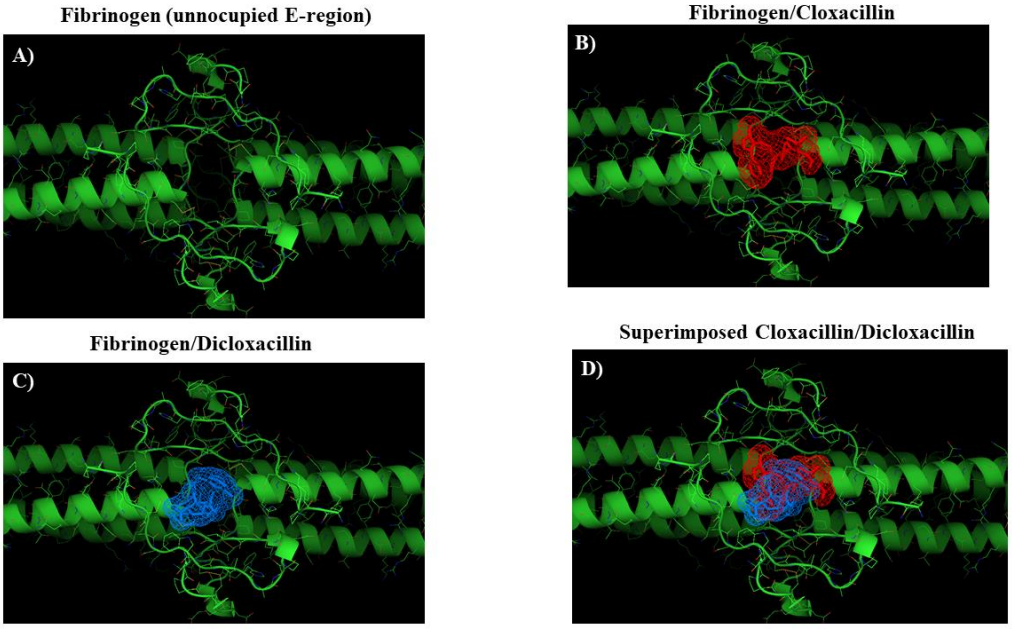


Figure S1. Pymol representation of the fibrinogen E-region binding site like: **A)** Unbound fibrinogen E-region, **B)** fibrinogen E region/cloxacillin complex with relative horizontal crystallographic orientation (red) respect to the binding site (green), **C)** fibrinogen E region/dicloxacillin complex with relative quasi-diagonal crystallographic orientation (blue) respect to the binding site (green), **D)** Superimposed cloxacillin and dicloxacillin docking complexes in the fibrinogen E-region binding site. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

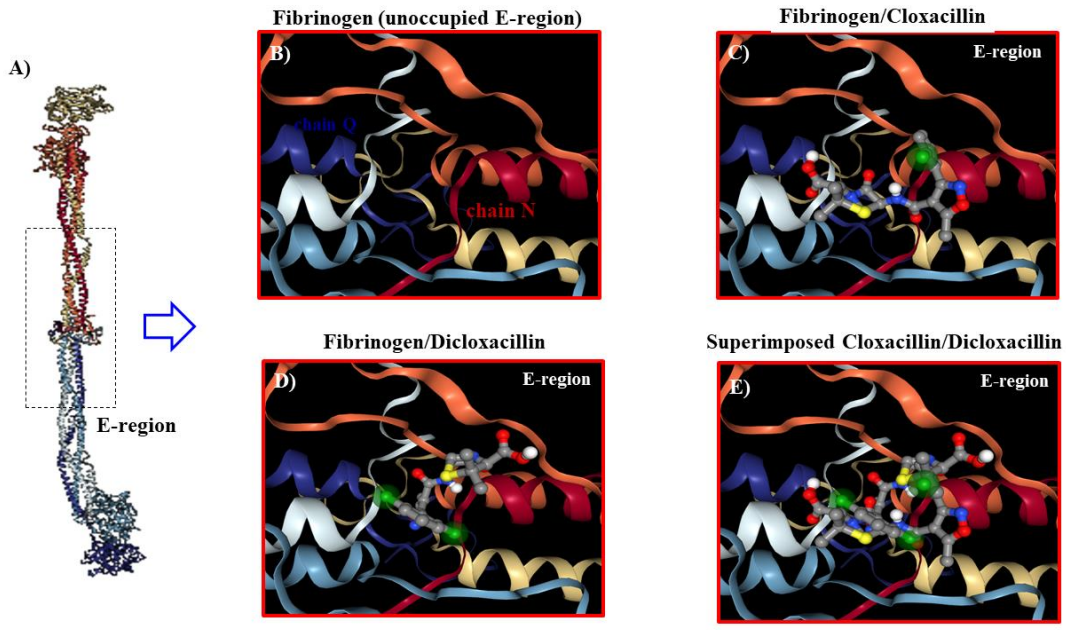


Figure S2. **A)** 3D-Representation of the whole molecular structure of the fibrinogen protein, **B)** representation of the apo-fibrinogen E-region (unbound sate) showing the two relevant target chain like chain N and chain Q marked as red and blue color-label, respectively. **C-D)** Best crystallographic docking complex obtained for both fibrinogen E-region/cloxacillin complex and fibrinogen E-region/dicloxacillin complex showing the different crystallographic orientation of the CL-atom(s) during the interactions highlighted in color green in the obtained 3D-lig-plot diagrams. Lastly, **E)** representation of the superimposed crystallographic docking complexes for both β-lactam antibiotics cloxacillin and dicloxacillin showing the relative position into the fibrinogen E-region binding site. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

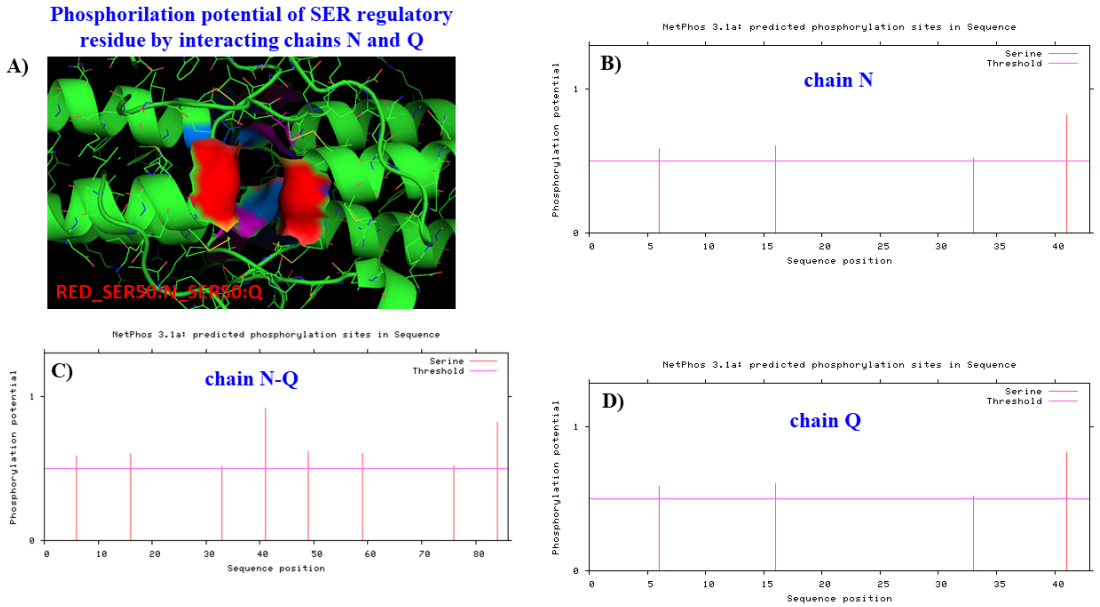


Figure S3. **A)** Representation of the fibrinogen E-region binding site (unbound state) highlighting as van der Waals surface (marked in red) the two key phosphorylation E-region regulatory residues (i.e., SER50:N and SER50:Q) which interact with both β-lactam antibiotics cloxacillin and dicloxacillin under study. **B-D**) Representation of the phosphorylation profiles for the fibrinogen E-region target chains (i.e., N, Q, and fused N-Q target chains). Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

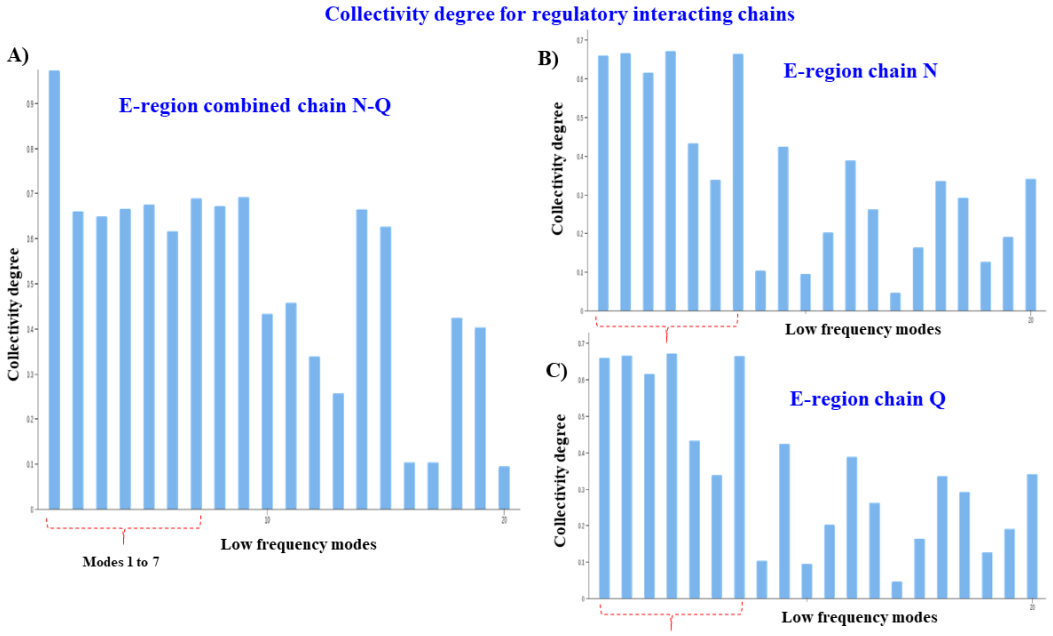


Figure S4. Graphical representation of the degree of collectivity (K_k) for the first seven low-frequency normal modes (k, from 1 to 7) obtained from the main fibrinogen E-region

target chains like **A)** fused N-Q chains, **B)** N chain, and **C)** Q chain using elastic network models. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

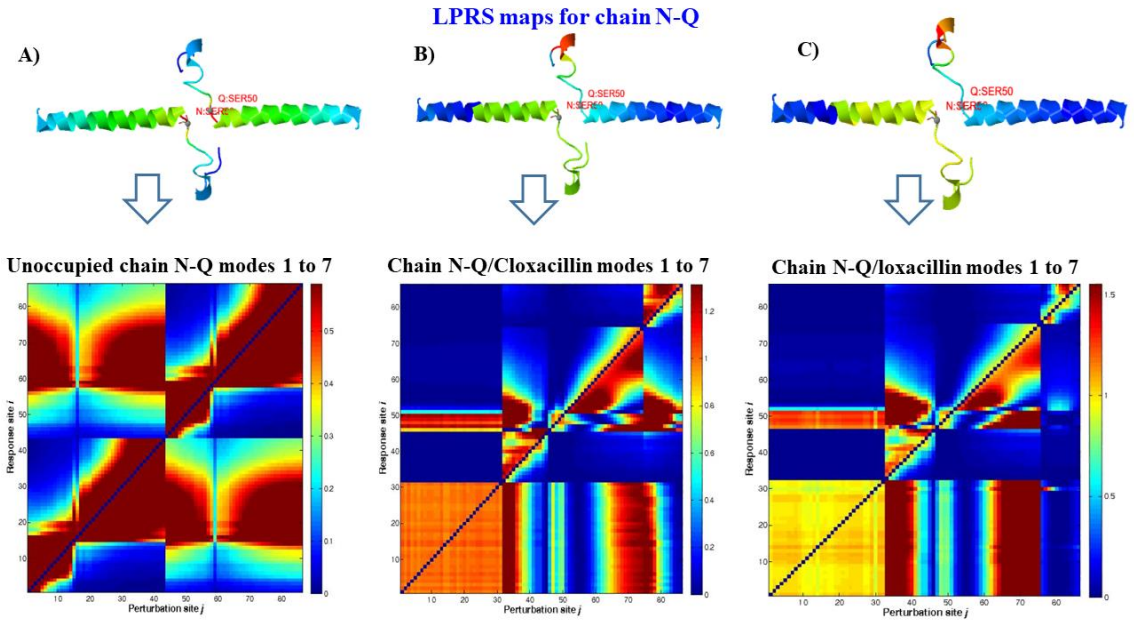


Figure S5. In the top, 3D-molecular structure of the fused N-Q target chains under different simulation conditions as **A)** unbound chain N-Q, **B)** chain N-Q plus cloxacillin, and **C)** chain N-Q plus dicloxacillin. In the bottom, is depicted the corresponding local perturbation response (LPRS maps) generated. Showing the relationship between (i-j)-pairs of residues as i-sensor residues (x-axis) to j-effector residues (y-ordinates) depicted like a 2D-matrix of allosteric network communication of inter-residue (i-j pairs as off-diagonal elements) and intra-residue (i-i pairs diagonal elements). On the far right of each LPRS map, the bar-color from blue to yellow regions correspond to weak to moderate local perturbations in the inter-residue communication for (i-j)-pairs of residues, and the orange to dark red regions represent strong inter-residue local perturbations for the aforementioned simulation conditions. All the simulations were performed for the first seven low-frequency normal modes (k, from 1 to 7). Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

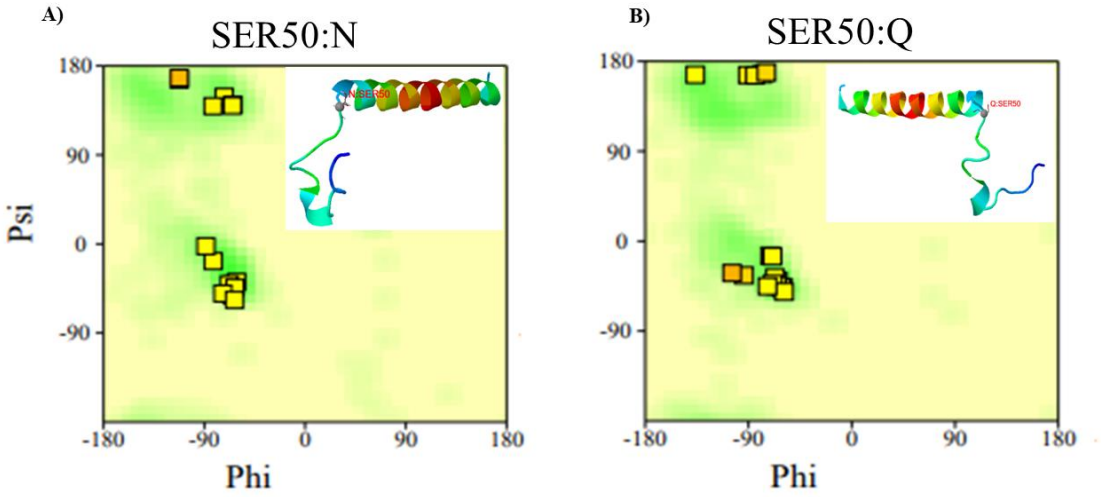


Figure S6. Validation-based Ramachandran diagrams for each phosphorylation regulatory residues as **A)** SER50: N and **B)** SER50: Q which belon to the fibrinogen E-region target chains (i.e., N and Q chains) and are directly involved in the docking interaction with the β-lactam antibiotics cloxacillin and dicloxacillin under study. Please note that, the evaluated target residues SER50:N and SER50:Q are conformationally-favored residues according to the presence in the light green contour into the individual Ramachandran diagram. Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

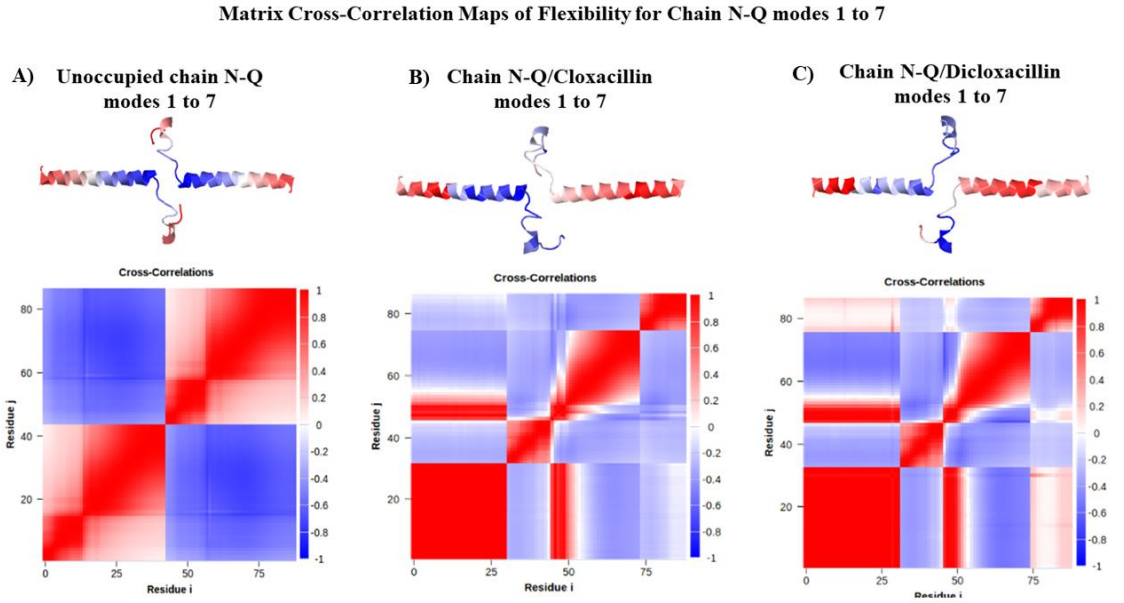


Figure S7. In the top, 3D-molecular structure of the fused N-Q fibrinogen E-region target chains under different simulation conditions as **A)** unbound chain N-Q, **B)** chain N-Q plus cloxacillin, and **C)** chain N-Q plus dicloxacillin. In the bottom, is represented the corresponding 2D-matrix-cross correlation maps of flexibility generated for the cited chains and plotted as a function of *i vs. j* residue indices in the absence and presence of the β-lactam antibiotics cloxacillin and dicloxacillin under study. For this instance, the *i*,*j*-pairs to fully correlated fluctuation-residues with strong correlations ($0.2 \leq C_{ij} \leq 1$) with color red-label pointing the same orientation for the *i* and *j* fibrinogen residues-flexibility. While the anti-correlated motion-based flexibility (*i.e.*: $C_{ij} \leq 0$) are colored in blue (*i.e.*, opposite-direction for the *i*,*j*-residue fluctuations and for moderately correlated and/or uncorrelated fibrinogen regions ($C_{ij} \approx 0$) are colored in light red and blue, respectively. Herein, all the simulations were performed for the first seven low-frequency normal modes (k, from 1 to 7). Source Elsevier article, please visit: https://doi.org/10.1016/j.molliq.2020.114667 (accessed on 15 february 2021)

Chapter 7.

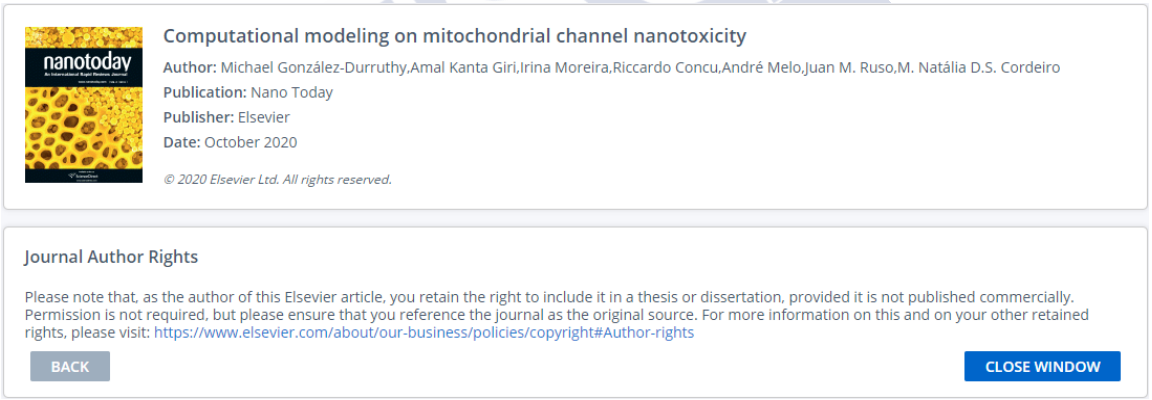
Michael González-Durruthy, Amal Kanta Giri, Irina Moreira, Riccardo Concu, André Melo, Juan M. Ruso, M. Natália D. S. Cordeiro. Computational Modeling on Mitochondrial Channel Nanotoxicity. *Nano Today*, 34, 100913, 2020. doi.org/10.1016/j.nantod.2020.100913.

Journal Impact Factor (JIF) (2020): 20.722.

CiteScore (2020): 20.9

Author Contribution M.G.-D.; Conceptualization, methodology-based on zig-zag-SWCNT modeling, molecular docking, flexibility properties evaluations of the hVDAC1, ENM-ANM approaches with perturbation maps, and writing—original draft preparation. *Corresponding author

Journal authorization:



Chapter 7.

Computational Modeling on Mitochondrial Channel Nanotoxicity

Abstract

Herein, we evaluated the interactions between the zig-zag-single-walled carbon nanotube (z-z-SWCNT (8.0)) and the ATP-entry-point of the human mitochondrial voltage-dependent anion-selective channel (hVDAC1). For this purpose, both molecular docking and molecular dynamics simulations were performed. The flexibility properties of the referred ATP-entry-point was efficiently modeled using crystallographic validation-based on Ramachandran plot. The preferred conformations obtained for this segment were able to establish very favorable interactions with the ligands (ATP and z-z-SWCNT). Next, using both molecular docking and molecular dynamics simulations, we demonstrated that z-z-SWCNT can directly prevent the ATP-transition from its first entry-point residue (MET1). We suggested that the associated z-z-SWCNT aggregation can be responsible by avoiding the natural biochemical steps for the ATP-transport, according to a nanotoxicity mechanism based on hydrophobic interactions. The docking free energy of z-z-SWCNT/hVDAC1 and ATP/hVDAC1 complexes was remarkably close, according to local perturbation maps of the catalytic residues' cluster (*i.e.* MET1, ARG2, GLY3, SER4, ALA5). On the other hand, the results of molecular dynamics simulations match the ones of the docking simulations, reinforcing the hVDAC1 channel nanotoxicity hypothesis. Overall, the obtained results could open new opportunities towards the rational design of new carbon nanomaterials and *in silico* mitotarget drug-discovery.

7.1 Introduction

The human mitochondrial voltage-dependent anion-selective channel (hVDAC1) is the main communication route between the mitochondrial inter-membrane space and cytosol.[1-3] From a structural point of view, the hVDAC1 channel is the most densely localized protein in the outer mitochondrial membrane of all human cells, which shows a β-barrel architecture composed by 19 β-strands with the N-terminal-(α)-helix segment located horizontally midway within the pore.[4-5] This segment adopts different

conformations in the voltage-gating channel, depending on the external and/or internal stimuli. In all the eukaryotic cells, the biochemical functions of this segment include for example: the ATP-bioenergetic regulation, membrane potential generation, the metabolic control cross-talk between the mitochondrial matrix and cytosol, the sucrose exchange, the intracellular [Ca^{2+}] regulation, the (Mg^{2+}, Zn^{+}, Na^{+}, K^{+}, etc.) ions transport, the regulation of glycolysis by association with hexokinase and the ADP/ATP-influx/efflux in hVDAC1.[5-6] In this regard, several pathological conditions-like: cancer diseases, Alzheimer, Parkinson's disease, epilepsy, acute cardiac ischemia and many others chronic processes have been associated with different degrees of hVDAC1-channel physiological perturbations leading to mitochondrial dysfunction by loss of regulation of the mitochondrial volume and promoting apoptosis.[5-7] Three specific regions, within the N-terminal-(α)-helix segment, are directly involved in the ATP-transport and efflux. These are: the ATP-entry point composed by the residue cluster (MET1, ARG2, GLY3, SER4, ALA5), the ATP-catalytic-point (ARG18) and the ATP-exit-point (LYS35). The most representative effect in the hVDAC1 channel dysfunction is the redox and bioenergetic imbalance, due to perturbations in the ATP-transport from the ATP-entry-point in the facing-inward conformation in mitochondrial matrix to the ATP-exit-point with facing outward conformation to the cytosol.[1, 8-10]

Nowadays, nanobiotechnology has brought a significant impact for many fields of modern science especially with the development of novel nanomaterials like carbon nanotubes.[11] Despite these, materials have been used for a wide range of biomedical applications; some evidences are also associating it with many hazardous effects.[12-15] This stresses the importance to develop appropriate methodologies (*in silico* tools), to predict and minimize the risk associated with those applications towards safe use.[12] The potential mitochondrial nanotoxicity (*i.e.,* bioenergetics dysfunction or mitotoxic effects) induced by single-walled carbon nanotubes (SWCNT), has been especially studied in the last 5 years due to their potential biomedical applications in mitochondrial nanomedicine for cancer therapy (*i.e.,* SWCNT inducing cytotoxicity-based mitochondrial bioenergetics dysfunction).[12-15] In this sense, new nanoinformatics approaches based on mitotarget modeling and drug-discovery concepts could offer a clear alternative to improve the studies on SWCNT with optimal structural nanodescritors for mitotropic behavior (mitochondrial selectivity), low off-target effects and good risk/benefit relationships. Currently, several scientific reports highlight that

SWCNT can induce potential mitochondrial channel nanotoxicity on critical human mitochondrial adenine-nucleotide transport channels like the ADP/ATP carrier and hVDAC1 channels.[12-15] According to this, previous computational-experimental studies on Nano-Quantitative Structure-Binding Relationships (QSBR) models[12-13] have suggested the existence of a singular correlation between the zig-zag geometry of the z-z-SWCNT (Hamada index m = 0, n > 0) with greater mitochondrial nanotoxicity-based reactivity compared with their non-zig-zag geometric counterparts like amchair-SWCNT (Hamada index m = n) and chiral-SWCNT (Hamada index $m \neq n$ and m > 0).[12-13] However, until the present, the mechanisms involved in the mitochondrial channel nanotoxicity-inducing ATP-bioenergetic dysfunction of carbon nanotubes remain largely ignored; and some computational methodologies could provide further insights on this subject. In particular, molecular docking and molecular dynamics (MD) simulations could be very useful frameworks for addressing these issues at an atomistic level.[12-18]

One of the main challenges, associated with these approaches, is the extremely high dimensionality of the ligand-protein conformational space. Particularly, the elastic normal mode with anisotropic network (ENM-ANM models) have been widely applied for describing the protein conformational dynamic, post-translational modifications, communication properties and signal propagation in the residue network of proteins under unperturbed (*i.e.*, native conformation in the absence of ligands) and/or perturbed conditions like induced by some given ligands (including nanoparticles).[19-22]

Herein, we suggest that the ENM-ANM approaches could be efficiently applied to capture the local-perturbations in the inter-residue network forming the ATP-entry-point of the N-terminal-(α)-helix segment of hVDAC1 channel, in the absence or presence of ligands (z-z-SWCNT and/or ATP the native hVDAC1-substrate).[22]

Herein, the main objective of this work is to investigate the ability of a zig-zag-single-walled carbon nanotubes (by simplicity referred as z-z-SWCNT (8.0) ligand-model) to induce mitochondrial channel nanotoxicity by selectively blocking the ATP-entry-point in hVDAC1, and disturbing the signal communication efficiency within its catalytic residue cluster (MET1, ARG2, GLY3, SER4, and ALA5). According to this, the immediate relevance of the modeling data could impact on: i) development of fast screening methods on potential human mitochondrial nanotoxicity using "human proteins" (*i.e.*, hVDAC1 crystallographic .pdb models) long before clinical studies in

human candidates, in accordance with ethical criteria in nanotoxicology, ii) efficient identification of therapeutically relevant channel-binding sites for mitochondrial nanomedicine, iii) rational optimization of structural determinants (nanodescriptors) of the channel mitotoxicity, and iv) improve the *in silico* methods on 3Rs principles (reduction, refinement and replacement) of animal testing in nanotoxicology research.

To probe the hypothesis, the critical interactions between the ligands (ATP native-substrate and z-z-SWCNT) and the ATP-entry-point of the hVDAC1 channel were evaluated. For this purpose, we used molecular docking, ENM-ANM analysis and molecular dynamics (MD) simulations.[12-18]

Finally, the present work gives a significant contribution to substantially improve the rational-design of new and safe carbon nanotubes. Open new opportunities for the promising mitochondrial nanomedicine, and making regulatory decisions in nanotoxicology.

7.2 Materials and methods

7.2.1 Theoretical Characterization of the Zig-Zag Single Walled Carbon Nanotube

For modeling and characterizing the electronic structure of the z-z-SWCNT crystal specie, the small crystal approach-based on the reciprocal zone-folding (ZF) method with Wannier orthogonal functions was used. [23-25] These latter functions can be expressed as follows:

$$\emptyset_{R-SWCNT}(r) = \frac{1}{\sqrt{N}}\sum_k e^{-ik\cdot R}\psi_k(r)\ ) \qquad \textbf{(1)}$$

where the Wannier functions, $\psi_k(r)$ represent the periodicity of z-z-SWCNT (8.0) ligand-model and are defined as:

$$\psi_k\,(r) = e^{ik\cdot r}u_k(r) \qquad \textbf{(2)}$$

In these equations, *r* represents the whole set of the atomic coordinates, and R is the lattice vector of the z-z-SWCNT structure, N corresponds to the number of cells in the z-z-SWCNT crystal in the reciprocal space that includes all the different *k*-values uniformly distributed in the Brillouin zone according to the applied periodic boundary

conditions.[23-25] The geometries of the ligands (z-z-SWCNT and ATP) were then optimized, using the semiempirical AM1 method as implemented in MOPAC.[24,23]

7.2.2 Molecular Docking

In this work, a molecular docking procedure was performed to ascertain the interactions between the ATP-entry-point in hVDAC1 and its ligands (ATP and z-z-SWCNT). For this purpose, two different complexes were studied: ***i***) hVDAC1 with the z-z-SWCNT ligand (z-z-SWCNT/hVDAC1), ***ii***) hVDAC1 with the ATP ligand (ATP/hVDAC1)

First, the X-ray structure of the receptor was retrieved from the *RCSB Protein Data Bank* (PDB).[23] This corresponds to the PDB entry 2jk4 with a resolution of 4.1 Å. This structure was then refined, by removing the water molecules and any other co-crystallized ligands. Additionally, hydrogen atoms were added to the hVDAC1 structure based on Gasteiger-Marsili empirical atomic partial charges, protonation states, bond orders and rotatable bonds.[26] To avoid obtaining false positives in the docking experiments, a hVDAC1 receptor structural validation was carried out through the Ramachandran maps of its ATP-entry-point residues.[33] This procedure allows verifying the absence of restricted flexibility for each of the hVDCA1 residues based on the allowed ψ *vs.* ϕ dihedral torsion angles of a given ATP-entry-point residue.

Before the docking runs, the potential hVDAC1-binding-sites including the ATP-entry-point were predicted using DeepSite.[28] This procedure was performed here by delimiting the van der Waals surface area of the ATP-entry-point binding residues, which bind to small ligands such as ATP or z-z-SWCNT. The DeepSite algorithm is based on a 3D-deep convolutional neural networks (DCNNs)[29-30] and was validated using more than 6500 proteins from the scPDB database.[27-28] The resulting ATP-entry-point docking box[32] was predicted to have a cubic form with size 16 Å and to be centered in the Cartesian point (X = 49.165 Å, Y = 13.086 Å, Z = – 6.072 Å).

All the molecular docking runs were performed using the AutoDock Vina package, with a Vina scoring function implemented by Trott and Olson.[16, 26]

Several docking runs starting from random orientations (or conformations) of the z-z-SWCNT ligand interacting with the ATP-entry-point of hVDAC1 were executed and ranked according to their stability in terms of the estimated free energy of binding.

In order to give reliable results, the AutoDock Vina exhaustiveness parameter equal to 100 (representing 100 z-z-SWCNT conformations) was established,[16, 26] which

determines the number of search to find best crystallographic ligand conformation in the docking complex formed with the ATP-entry-point from "human protein" (*i.e.*, hVDAC1 crystallographic .pdb model).

The docking results were found as energetically unfavorable when the FEB values were lower or equal than 0 kcal/mol (worst crystallographic poses), clearly denoting either an extremely low or complete absence of affinity related to a repulsive pattern of interactions. Notice that the FEB of such complexes was computed considering a cutoff value of 7 Å for the interatomic distances $d_{i,j}$ between the *i*-ligands (z-z-SWCNT or ATP) and *j*-target residues of the ATP-entry-point, *i.e.*, only atoms with distances $d_{ij} \leq 7$ Å were considered as interacting atoms.[29-31] Moreover, the best root-mean-square deviation (RMSD; eq **3**) was stored as a criterion of correct docking pose accuracy for atomic positions below 2 Å.[29-30, 32-34]

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n_L}\sum_{j=1}^{n_P}\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i - z_j\right)^2}{n_L \times n_P}} \qquad \textbf{(3)}$$

In equation (3), n_L is the number of atoms of the ligand and n_P is the number of atoms of the hVDAC1 protein. Additionally, each term included in the double summation represents the distance between a ligand atom *i* with Cartesian coordinates (x_i, y_i, z_i) and a protein atom *j* with Cartesian coordinates (x_j, y_j, z_j).

7.2.3 ENM-ANM approach.

The elastic normal mode with anisotropic network models (ENM-ANM) is a computational approach which evaluates the degree of perturbation (strength of interactions), induced by a given ligand (z-z-SWCNT or ATP) in the residues network of a given protein (ATP-entry-point of hVDAC1), by describing the interaction potential (*V*) for a receptor-ligand complex as a Hookean potential (or 'springs') based on elastic normal mode analysis.[19-22] The mathematical framework of this approach has been extensively reported, and thus we will highlight only the more important aspects.

The ANM approach was performed, for each C(α)-atom of the N-terminal-(α)-helix residues, in the absence or presence of ligands.[18-22] Two VDAC1-residues were assumed

to be in contact when their C(α)-atoms are separated by less than a cutoff distance (Rc ≈ 7 Å), which is directly related to effective communication distances ($^{eff}S_{(ij)}$ < 3,81 Å) between C(α)-atoms of consecutively connected (*i*,*j*)-residues. The latter are identified by the mean-square distance traveled by a random walk of *n*-Markov transition steps according to the following equation,

$$(^{eff}S_{(ij)})^2 = n(L)^2, \quad \textbf{(4)}$$

where *L* is the average step size.

The N-terminal α-helix residues were then represented as a transition matrix network that is controlled by Markov transition probabilities for the passage of information across the nodes (C(α)-atoms-hVDAC1-residues) to express the intrinsic fluctuation in the N-terminal-(α)-helix segment.[17]

The influence of the ligands (z-z-SWCNT and ATP) on the hVDAC1 inter-residue communication (N-terminal-(α)-helix segment) was evaluated by applying a local-perturbation response scanning (LPRS) analysis, where the perturbation strength of each ATP-entry-point residue (*i.e.* MET1, ARG2, GLY3, SER4, ALA5) has over its neighbor residue is based on correlated and/or uncorrelated motion relationships.[19-22]

The collective fluctuation of the anisotropic network is completely defined by a 3*N* × 3*N* Hessian matrix (*H*),

$$H_{i,j} = \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,N} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,N} \\ . & . & \cdots & . \\ H_{N,1} & H_{N,2} & \cdots & H_{N,N} \end{bmatrix} \quad \textbf{(5)}$$

in which its *N*-elements ($1 \leq i, j \leq N$) are the second derivatives of the anisotropic network potential (*V*):

$$H_{i,j} = \begin{bmatrix} \partial^2 V / \partial X_i \partial X_j & \partial^2 V / \partial X_i \partial Y_j & \partial^2 V / \partial X_i \partial Z_j \\ \partial^2 V / \partial Y_i X_j & \partial^2 V / \partial Y_i \partial Y_j & \partial^2 V / \partial Y_i \partial Z_j \\ \partial^2 V / \partial Z_i X_j & \partial^2 V / \partial Z_i \partial Y_j & \partial^2 V / \partial Z_i \partial Z_j \end{bmatrix} \quad \textbf{(6)}$$

On the other hand, the 3*N*-dimensional vector $\Delta R(i)$ represents the residue displacements of hVDAC1 in *response* to the application of a *perturbation* induced by the ligand (*i.e.*: following Hooke's law), and can be represented by,

$$\Delta R(i) = H^{-1} F^{(i)}, \quad \textbf{(7)}$$

where the strength of perturbation ($F^{(i)}$ exerted on residue (i) is expressed according to:

$$F^{(i)}{}_{1\times 3N} = \{0 \quad 0 \quad 0 \quad \cdots \quad F_x^{(i)} \quad F_y^{(i)} \quad F_z^{(i)} \quad \cdots \quad 0 \quad 0 \quad 0\}^{\mathrm{T}} \quad \textbf{(8)}$$

Once it has been set up, it is possible to obtain the inverse of the H-matrix ($H_{i,j}{}^{-1}$) according to the following equation:

$$H = \frac{1}{H_{i,j}} = -\frac{1}{s_{i,j}^{p+2}} \begin{bmatrix} (X_j - X_i)(X_j - X_i) & (X_j - X_i)(Y_j - Y_i) & (X_j - X_i)(Z_j - Z_i) \\ (Y_j - Y_i)(X_j - X_i) & (Y_j - Y_i)(Y_j - Y_i) & (Y_j - Y_i)(Z_j - Z_i) \\ (Z_j - Z_i)(X_j - X_i) & (Z_j - Z_i)(Y_j - Y_i) & (Z_j - Z_i)(Z_j - Z_i) \end{bmatrix} \quad \textbf{(9)}$$

($H_{i,j}{}^{-1}$) may be used to predict the auto- and cross-correlations of residues, and it represents the covariance matrix of *3N* multi-variant distribution. This matrix holds the relevant information on the hVDAC1-residues fluctuations (*i.e.*, the residues from ATP-entry-point involved in the formation of stable ligand/hVDAC1 complexes); $s_{i,j}$ represents the instantaneous distance between nodes or residues (i and j); and p is an empirical parameter.[19-22]

The main idea of the proposed LPRS analysis is to quantify the force of the perturbations induced by the ligand (z-z-SWCNT) in the ATP-entry-point target-residues in order to determine the local-perturbation response in the overall residue-network following **eq 7**. As mentioned above, $\Delta R\ (i)$ is a *3N*-dimensional vector that describes the displacements of all the residues (*i.e.*, in *N*-blocks of dimension 3, each) under perturbation $F^{(i)}$.[22] The response of a given *k*-residue is the magnitude $\left\langle \left\| \mathbf{\Delta R}_k^{(i)} \right\|^2 \right\rangle$ of the k^{th} block of $\Delta R\,(i)$ averaged over multiple $F^{(i)}$, expressed as an element of the $3N\times 3N$ LPRS matrix.[22]

The response to local displacements at each perturbation site is obtained by dividing each row by its diagonal value (d) as shown in the following equation,

$$\overline{\boldsymbol{S}}_{\mathrm{LPRS}} = \begin{pmatrix} 1/d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1/d_N \end{pmatrix} \mathrm{LPRS} \quad \textbf{(10)}$$

where each i^{th} row of $\bar{S}_{LPRS}$ refers to the profile generated by the perturbing residue (i).

7.2.4 Molecular Dynamics Simulations.

To understand the interactions between ligands and hVDAC1 (ATP-entry-point) at nanometric level, we performed MD simulations for three different systems as i) ATP + hVDAC1, ii) z-z-SWCNT + hVDAC1, and iii) co-interacting system z-z-SWCNT+ATP/hVDAC1. The initial structures of the ligand-hVDAC1 were taken from the molecular docking results and placed each of them at the center of a three-dimensional periodic cubic box with size 8.9 nm x 8.9 nm x 8.9 nm. Afterwards, the boxes were filled with 21,910 water molecules, and few Cl^- ions were added to neutralize the systems.

The GROMACS-5.1.4 software package[35-36] was used for performing all the MD simulations. The intra- and inter-molecular interactions of the ligand-hVDAC1 protein were modeled by the GROMOS 43A1 force field parameter set, [37-38] and the water molecules by the SPC/E potential[39] with rigid bonds and angle. The cross interactions between unlike atoms were calculated by applying geometric combination rules.[40-42] The steepest descent algorithms were used for 50000 steps to optimize the initial geometry keeping the ligand-hVDAC1 complex rigid.

Thereafter, systems were equilibrated in two steps: **1)** the systems were simulated for 100 ps in NVT ensemble (number of atoms volume and temperature constant) to stabilize the system's temperature at 300 K, **2)** the systems were then equilibrated for 100 ps in NPT ensemble (number of atoms, pressure and temperature constant) to stagnate the pressure of the systems at 1 bar. After these equilibration steps, MD simulations were performed for 20 ns in *NPT* ensemble to collect data for analysis. This production run was considered appropriate, once it was observed the distance between the center of mass of the ATP-entry-point and the center of mass of the ligand's constant for first 6 ns of the production run. The leap-frog algorithm[43], with a time step of 2 fs, was employed to integrate the equation of motion for the systems. The system pressure was controlled by a Parrinello-Rahman method[44] with a time constant of 2.0 ps, and the system temperature was kept close to the intended value of 300 K, using the v-rescale coupling algorithm[45] with a time constant of 0.1 ps. The cut off distance for both the Lennard-Jones potential and short-range Coulombic interactions were 1 nm.[46] The long-

range Coulombic interactions were calculated by the Particle Mesh Ewald (PME) method.[47-48]

The binding free energy of the complexes was calculated using the Molecular Mechanics Poisson Boltzmann Surface Area (g_mmpbsa) method implemented in GROMACS 5.1.4.[49] The binding free energy of the complexes was analyzed by taking trajectory frames at an interval of 1 ps of the simulations. The binding free energy can be expressed as follows,

$$FEB \approx \Delta G_{bind} = G_{complex} - \left(G_{hVDAC1} + G_{ligand}\right) \quad \textbf{(11)}$$

where $G_{complex}$ is the total free energy of the ligand/hVDAC1 (ATP-hVDAC1, z-z-SWCNT-hVDAC1 or ATP-z-z-SWCNT/hVDAC1) complex, and G_{hVDAC1} and G_{ligand} are the total free energy corresponding to the separately solvated hVDAC1 and ligand (ATP, z-z-SWCNT or ATP-z-z-SWCNT), respectively. The free energy ($G(x)$) for each individual species x (ligand, hVDAC1 or ligand/hVDAC1 complex) can be expressed as depicted in equations **11** and **12**.

$$G(x) = \langle E_{MM}(x) \rangle - T\,S(x) + \langle G_{solvation}(x) \rangle \quad \textbf{(12)}$$

$$E_{MM}(x) = E_{bonded}(x) + E_{nonbonded}(x) = E_{bonded}(x) + E_{vdw}(x) + E_{elec}(x) \quad \textbf{(13)}$$

In equation (13), $\langle E_{MM}(x) \rangle$ is the average molecular mechanics potential energy of the molecular specie x, $S(x)$ is the respective entropy and $\langle G_{solvation}(x) \rangle$ is the average value of its solvation free energy. According to equation (8), the potential energy ($E_{MM}(x)$) can be expressed as the sum of its bonded ($E_{bonded}(x)$), van der Waals ($E_{vdw}(x)$) and electrostatic ($E_{elec}(x)$) components. The solvation free energy ($G_{solvation}(x)$) is expressed as the sum of the polar ($G_{polar}(x)$) and nonpolar ($G_{apolar}(x)$) solvation free energies according to the following equation,

$$G_{solvation}(x) = G_{polar}(x) + G_{apolar}(x) \quad \textbf{(14)}$$

where $G_{polar}(x)$ is estimated solving the Poisson-Boltzmann equation and $G_{apolar}(x)$ is calculated from the solvent accessible surface area ($SASA(x)$) using the following expression,

$$G_{apolar}(x) = \gamma SASA(x) + b \quad \textbf{(15)}$$

where γ and *b* are the empirical constants.

7.3 Results and discussion

7.3.1 Ligand selection, characterization and modeling of z-z-SWCNT.

Let us first start by focusing on the well-known direct proportionality relationship between the carbon nanotube zig-zag geometry with the nanotoxicity-based reactivity at the subcellular level (mitochondria).[12-13] The presence of dangling bonds in the tip of the z-z-SWCNT has a direct influence on the modeled biochemical response (*i.e.*, mitochondrial channel nanotoxicity). Herein, we select the zig-zag topology (Hamada index m = 0, n > 0) like a robust and recognized electro-topological nanodescriptor of single-walled carbon nanotube ligand (z-z-SWCNT). Because, it is well known that the potential mitochondrial channel nanotoxicity properties of z-z-SWCNT are greater compared to its non-zig-zag geometric counterparts as amchair-SWCNT (Hamada index m = n) and chiral-SWCNT, when it has Hamada index (n, m), with m > 0 and m ≠ n; and its enantiomer (or mirror image) has Hamada index (m,n), which is different from (n, m) no reflection symmetry.[12-13]

As reported in previous studies, a unique phenomenon of "edge effects" present only in semiconducting pristine z-z-SWCNT geometries has paramount importance like structural nanodescriptor of mitochondrial channel nanotoxicity.[12-13] Indeed, the presence of tip-charge variations in the open z-z-SWCNT-extremes from the cycloparaphenylene aromatic system; increases the reactivity of these carbon nanotubes when compared with their non-zig-zag geometric counterparts like armchair-SWCNT and chiral-SWCNT.[12-13] Besides, the influence of dangling bonds in the z-z-SWCNT tested, induce appreciable changes in the structural and electronic properties of the z-z-SWCNT system, partially mimicking the effects of topological vacancies observed in the SWCNTs with side-wall defects. It is well known that the z-z-SWCNT dangling bonds can influence the Fermi energy. In fact, the existence of more sp^3-C-atoms can induce a flat-band magnetism (magnetic polarization) in the Fermi level, that together with the π-electrons produce a larger increase in the intrinsic reactivity of the z-z-SWCNT topologies, maximizing the nanointeractions and consequently its mitochondrial channel nanotoxicity (*i.e.*, more negative expected FEB values obtained

from docking and molecular dynamic simulations) compared its non-zig-zag geometric counterparts (amchair-SWCNT and chiral-SWCNT).

In this context, it is important to note that during biochemical interactions of a given pristine-SWCNT with mitochondrial redox-components, dangling bonds formation from non-zig-zag geometric counterparts (amchair-SWCNT and chiral-SWCNT) can also be spontaneously generated at the subcellular level (mitochondria).

Then, we performed the ligand selection, modeling and characterization of the electronic structure using a validated z-z-SWCNT (8.0) ligand model to ensure the quality of our nanotoxicological predictions on hVDAC1 (ATP-entry-point).[12-13]

Herein, we present the results of the modeled z-z-SWCNT (8.0) ligand used in our study, which was obtained by using the reciprocal ZF scheme introduced by Saito and Dresselhaus,[24-25] which is considered a good approximation for 2D Brillouin zone of graphene, within the tight-binding electronic band structure approach when applying Wannier orthogonal functions (see **Figure 1**).[23-25]

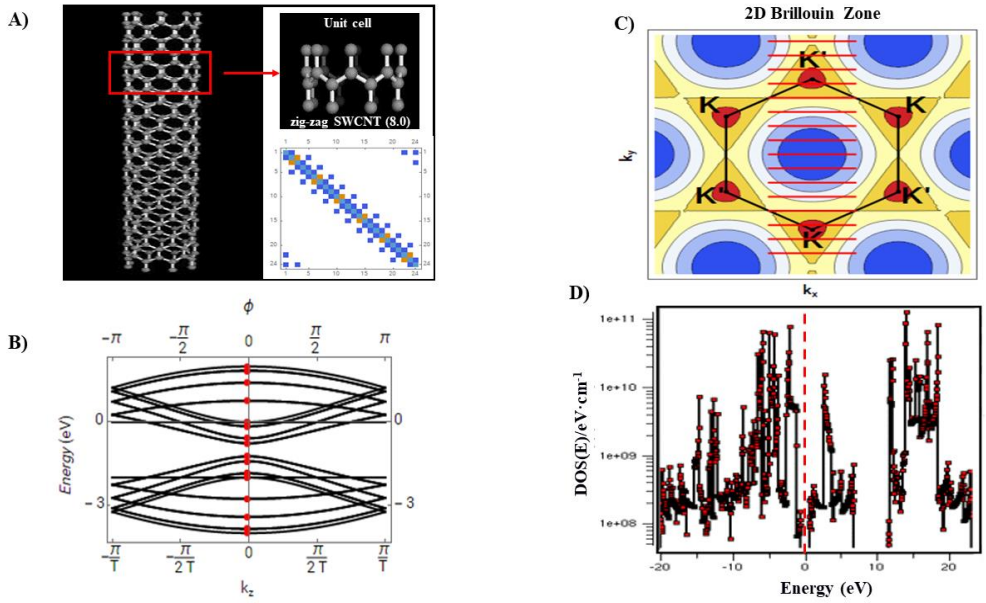


Figure 1. Electronic structure z-z-SWCNT obtained using the small crystal approach based on the reciprocal zone-folding (ZF) method with Wannier orthogonal functions. **A)** Molecular model used, with the corresponding unit cell showing an alternative way to depict the reciprocal space zone-folding (ZF) like a tight-binding Wannier representation. **B)** Electronic band structure representation of Energy (eV) *vs*. k_z,

performed by ZF method setting the following tight-binding parameters: hopping parameter $t = -1$, on-site energy parameter $\epsilon = -1$ and overlap $s = 0$. **C)** Contour plot, for the 2D Brillouin zone, obtained by applying the ZF method in the 2D-graphitic Brillouin zone with Dirac points (K and K′) uniformly distributed (red cutting lines) in the (k_x, k_y)-plane. **D)** Representation of the density of electronic states (DOS) profile as a function of the energy (E) counting all over the states within each interval [E, $E + dE$]. Source Elsevier article, please visit: https:// doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

7.3.2 Molecular docking results. A key step to ensure the accuracy of molecular docking data involves the correct prediction of suitable ATP/hVDAC1 binding sites. Many algorithms for detecting protein-binding cavities have been settled over the years based on the crystallographic structure, geometric, and chemical protein 3D-features (ATP/hVDAC1). In the present study, the prediction of the binding residues from hVDAC1 (the ATP-entry-point) was carried out using the machine learning algorithm based on 3D-deep convolutional neural networks (DeepSite-CNNs chemoinformatic tool), [50] which was previously validated by an extensive test set from the scPDB database (> 7000 proteins structures).[51-52] The structural identification results thereby obtained for the hVDAC1 binding-sites including the ATP-entry-point are presented in **Figure 2**.

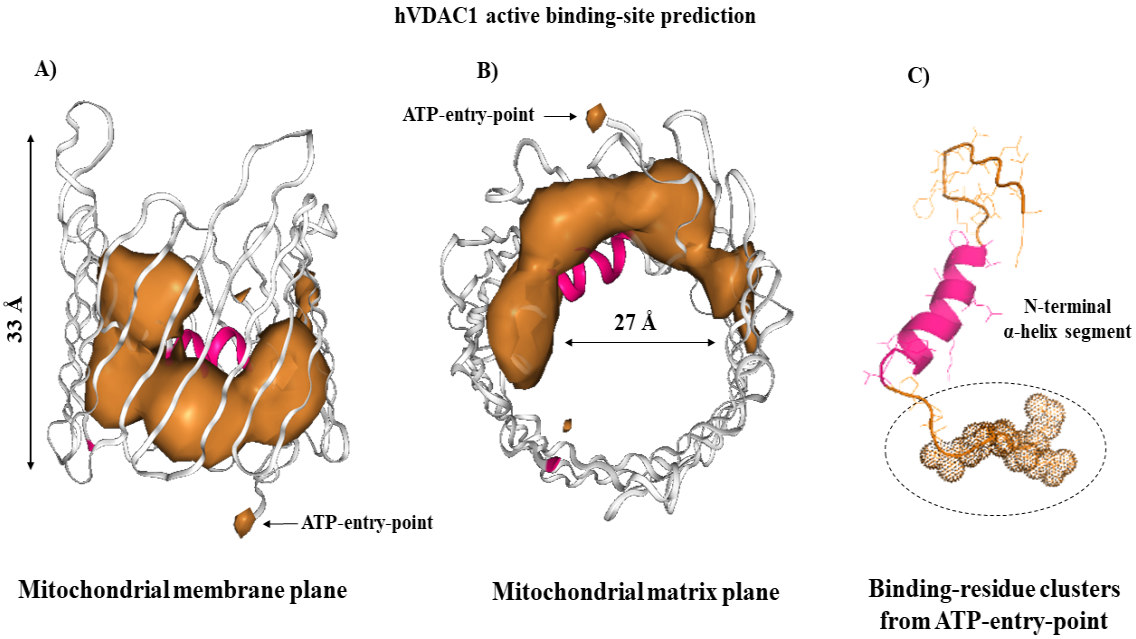


Figure 2. **A)** DeepSite prediction of topological cavities of the hVDAC1 active binding sites (volumetric orange regions) from lateral view. **B)** Display of the hVDAC1active binding-sites (volumetric orange regions) from the mitochondrial matrix plane showing the ATP-entry-point. **C)** Representation of the N-terminal-(α)-helix segment including the binding-residue clusters (MET1, ARG2, GLY3, SER4, ALA5) of the hVDAC1 channel. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

Following on, to validate the 3D X-ray crystallographic structure of the N-terminal α-helix segment a Ramachandran analysis was carried out. This analysis comprised plotting a 2D-projection on the plane of the 3D-crystallographic structure of protein hVDAC1, describing all possible conformations of the ATP-entry-point residues by considering the dihedral angles (ψ and ϕ) around the peptide-bond of such residues.[53] The allowed ψ *versus* ϕ torsion values of a given residue of hVDAC1, found within the Ramachandran purple colored contour, were considered as conformationally favored residues. Otherwise, those were considered as sterically disallowed residues (see **Figure 3**). As seen, the Ramachandran plot obtained corroborates the very good X-ray crystallographic quality of the N-terminal-(α)-helix segment model with an expected

percentage > 90% for favored regions. It is also important to note that, the LYS23 and ALA11 (Ramachandran outliers) are not part of the ATP-entry-point.

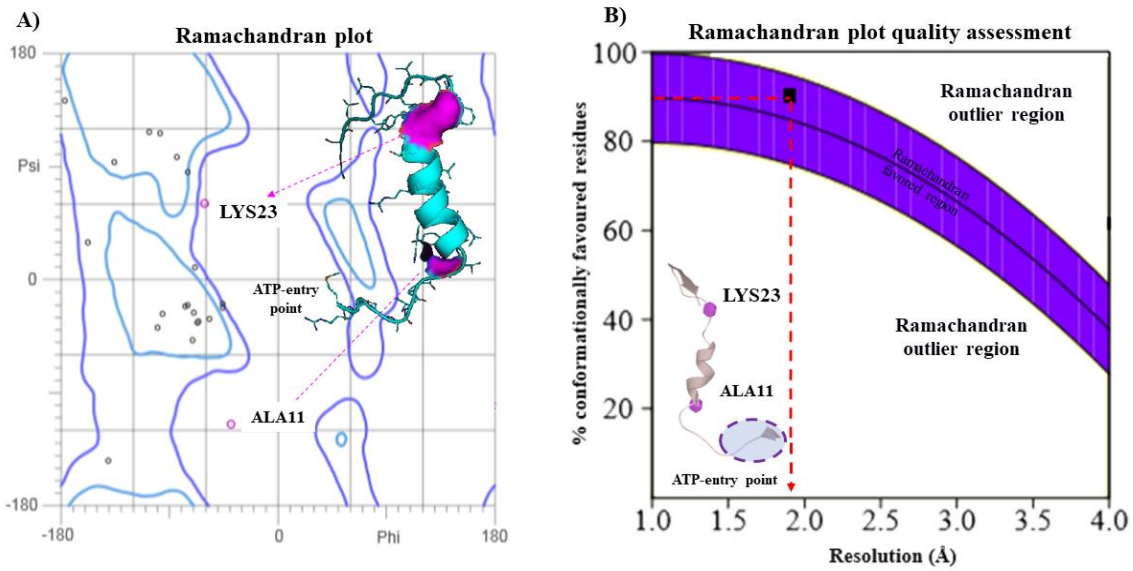


Figure 3. **A)** Ramachandran plot and spatial distribution of its outliers (light purple residues: ALA11 and LYS23) only for the N-terminal α-helix segment of the hVDAC1-channel PDB X-ray structure. All the possible combinations of Psi *vs.* Phi dihedral torsion angles of each amino acid residue of the hVDAC1-channel protein are shown. **B)** Ramachandran plot quality assessment (hVDAC1-channel model quality) measured by the percentage of the hVDAC1-residues that are in the most favored residues localized in the purple shaded region. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

To elucidate the critical regions that could be affected by z-z-SWCNT-protein interactions, we evaluated also the flexibility properties of the hVDAC1 general structures considering the N-terminal-(α)-helix domain.[1-9] Particularly, it can be seen in **Figure 4** that the ATP-entry-point's flexibility properties are crucial to investigate the mitochondrial channel nanotoxicity-based bioenergetic dysfunction.

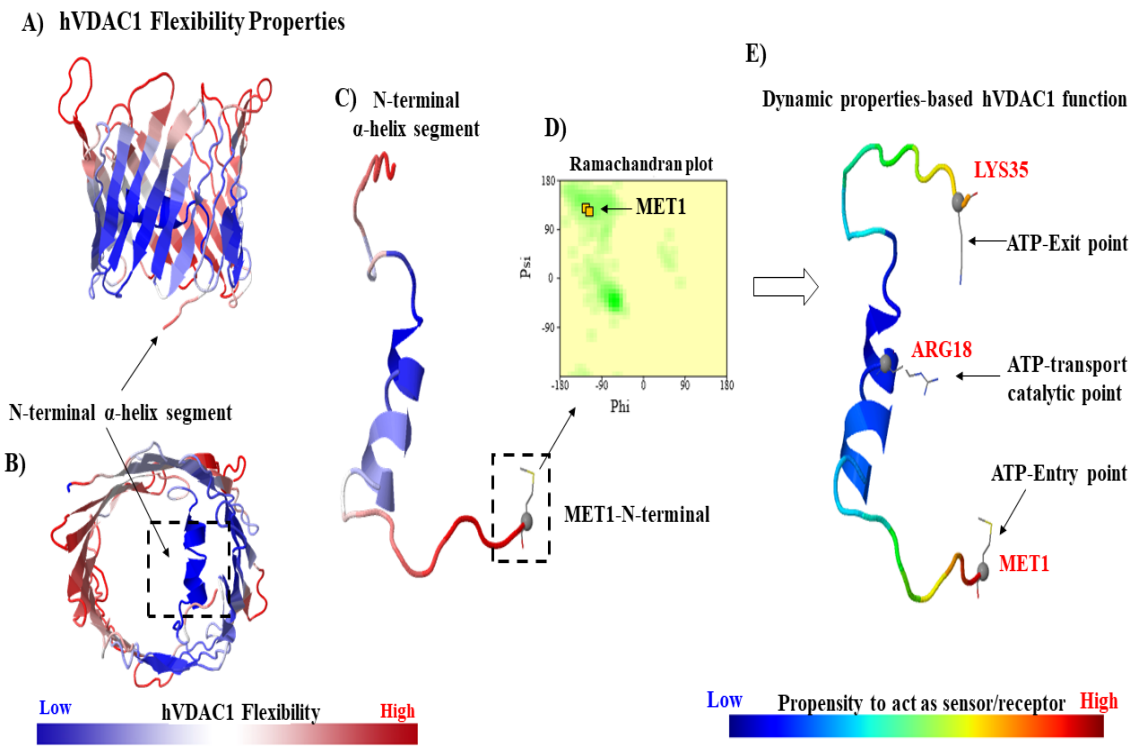


Figure 4. Schematic representation of the hVDAC1 flexibility properties. The colored structures are based on the size of fluctuations of the hVDAC1-residues showing low-flexibility (blue) to high-flexibility (red). **A)** Lateral view of hVDAC1. **B)** Mitochondrial matrix plane view of hVDAC1. **C)** The flexibility properties of N-terminal-(α)-helix segment including the MET1-flexibility corresponding to the first residue of the ATP-entry-point. **D)** Ramachandran validation based on ψ *vs.* ϕ dihedral torsion angles for the first ATP-entry-point residue (MET1) with allowed conformational properties (inner light green region). **E)** Representation of the N-terminal-(α)-helix based on hVDAC1 dynamic functional properties. Herein, it shows the behavior (like effector or sensor) of the most critical residues of the ATP-entry-point in the absence of z-z-SWCNT, such as MET1: effector-residue (red region), ATP-catalytic point ARG18 residue: sensor-residue (blue region), and the ATP-exit point LYS35 residue: chain A receptor-behavior (red region). Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

After this, we carried out the molecular docking experiments to obtain the Gibbs free energy of binding (FEB) for the three complexes under study. The results obtained show that the FEB values are very similar for the two binary complexes (–3.7 kcal·mol^{-1}

for z-z-SWCNT/hVDAC1 and –3.8 kcal·mol^{-1} for ATP/hVDAC1). These results demonstrate the reliability of the proposed mechanism for the z-z-SWCNT channel-nanotoxicity, based on hydrophobic and electrostatic interactions established between this ligand and the ATP-entry-point residues.

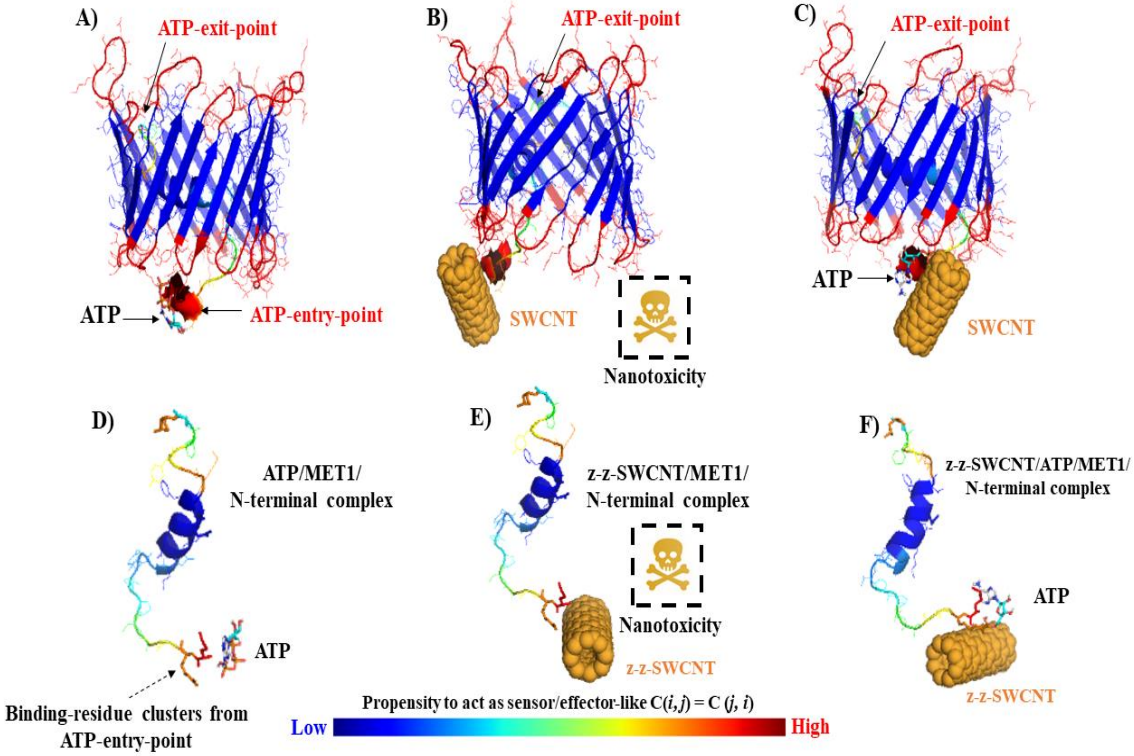


Figure 5. On the top row, snapshots showing the molecular docking results for the best binding-poses like: **A)** docking control ATP/hVDAC1 complex (–3.8 kcal/mol), **B)** z-z-SWCNT/hVDAC1 complex (–3.7 kcal/mol) and **C)** overlapping representation of the best binding-poses of ATP and the z-z-SWCNT at the ATP-entry-point. The ATP-entry-point residues (MET1, ARG2, GLY3, SER4, and ALA5) are depicted as small van der Waals surfaces labeled-red. In the bottom row, the docking complexes like **D)** ATP/hVDAC1, **E)** SWCNT/hVDAC1 and **F)** overlapping binding-poses of ATP and SWCNT are presented considering the best binding-poses of these ligands interacting simultaneously with the ATP-entry-point from the N-terminal-(α)-helix segment. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

The docking results, depicted in **Figure 5**, suggest that the interactions between z-z-SWCNT and the ATP-entry-point residues are hydrophobic. Furthermore, it is important

to highlight that the regulatory SER-(4)-N-terminal α-helix residues have an important role in the phosphorylation mechanism for controlling the function of ATP-entry. The presence of z-z-SWCNT could significantly weaken the physiological function pathways and induce post-translational modifications in SER4.[12, 18] It is well-known that hVDAC1 post-translational modifications are associated with biochemical reactions of phosphorylation. According to this, we thus suggest that the z-z-SWCNT might mimic pathological conditions like post-translational modifications in the SER4 of the ATP-entry-point such as metabolic stress, aging, cancer, and cardiovascular diseases.[6]

Furthermore, it resembles that z-z-SWCNT initially blocks the ATP-transition from the first residue (MET1) of the ATP-entry-point before it interacts with the positively charged ε-amino groups (N^+ primary amines) of ARG2-hVDAC1. The amino groups (N^+) of ARG2-hVDAC1 are highly reactive and often participate in reaction at the ATP-entry-point acting as a voltage sensor in mitochondria.[4, 18] Nevertheless, the presence of z-z-SWCNT and ATP molecules in the same biophysical environment could induce local-perturbations in the inter-residue communication, and the transduction properties of the ATP-entry-point. This would affect the signal communication efficiency in the N-terminal-(α)-helix segment and the subsequent steps of the ATP-translocation through hVDAC1 like ATP-transport in the catalytic point (ARG18) and the ATP-efflux point (LYS35).[7-9] Overall this thus suggest that the molecular docking mechanism is likely-based on local-perturbations (interactions) with effector and allosteric residues from the ATP-entry-point.[19]

To further evaluate the local perturbation on ATP-entry-point residues in the absence or presence of ligands (z-z-SWCNT and ATP-native substrate of VDAC), we applied an anisotropic network approach based on the biophysical parameter Markov commute-time [C (i, j) →C (j, i)].[19,54] The latter represents the ability of a given target to receive (sensor residues) or send (effector residues) perturbation signals in both directions ($j \rightarrow i$ and $i \rightarrow j$). See **Figure 6**.

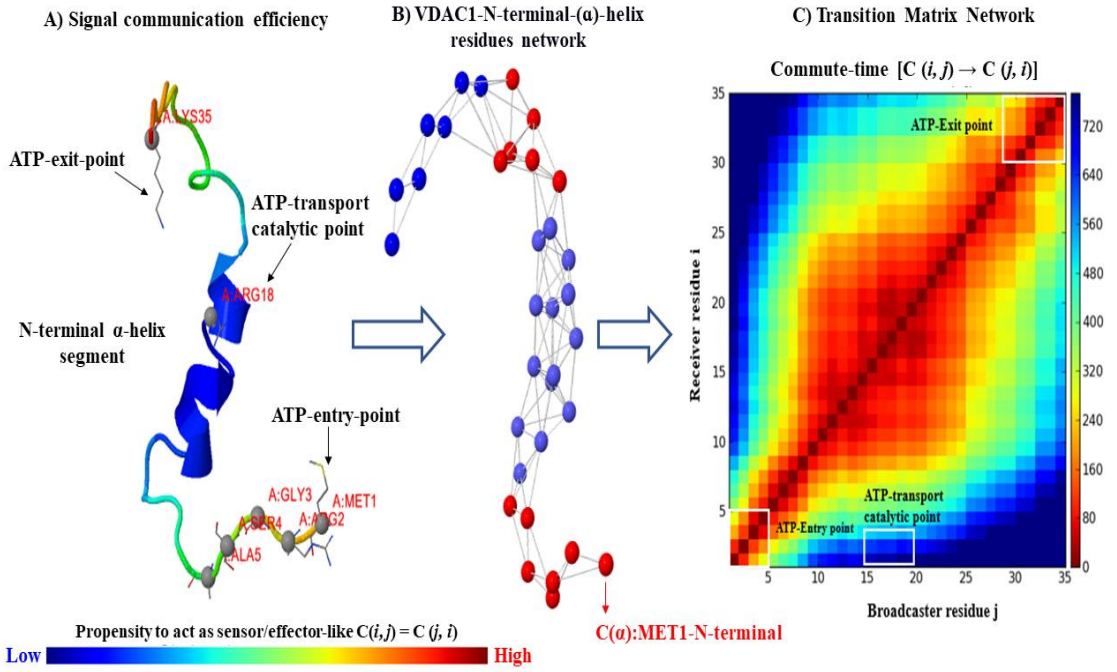


Figure 6. **A)** Representation of intra-molecular communications between residue's fluctuation in the N-terminal-(α)-helix segment based on the efficiency of sensors (blue) and effectors (red): ***i***) ATP-entry-point ***ii***) ATP-transport in the ATP-catalytic-point (ARG18) ***iii***) ATP-exit-point (LYS35) **B)** N-terminal-(α)-helix residues network of C(α)-atoms connected by elastic springs. Herein, the color of atoms (spheres) is depicted depending on the B-factor. **C)** Two-dimensional transition matrix network-based on the Markov commute time (C (*i*, *j*)), clustering sensor and effector residues in the N-terminal-(α)-helix. The critical regions of ATP-transport stages are represented by white rectangles. The color bar on the right side indicates low (blue) and high (red) commute time. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

Following this idea, we theoretically modeled the local-perturbation response induced by the ligands in the ATP-entry-point. The results of this local-perturbation response scanning analysis are presented in **Figure 7**.

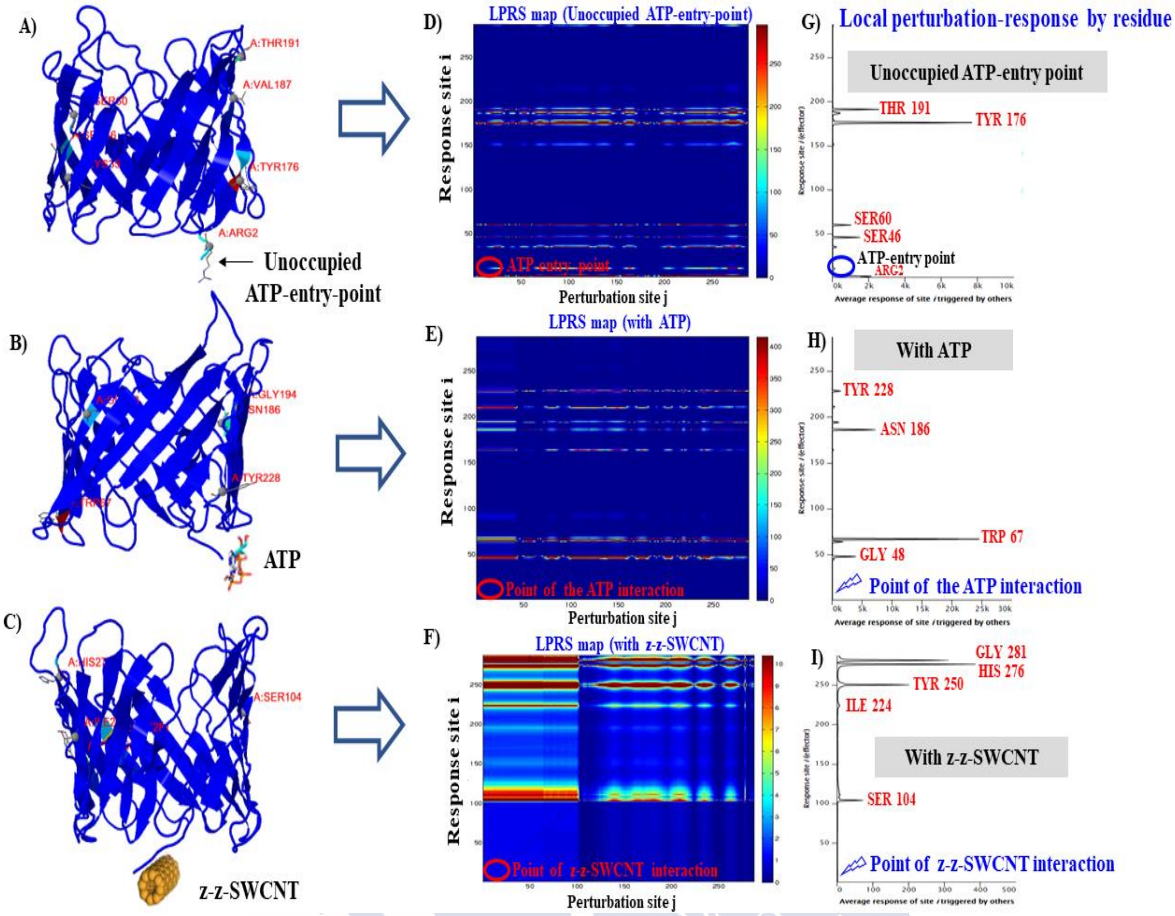


Figure 7. On the left, 3D-cartoon representations for the simulated conditions, namely: **A)** hVDAC1 without ligand, **B)** ATP/hVDAC1 complex, and **C)** z-z-SWCNT/hVDAC1 complex. On the right, the corresponding local perturbation-response scanning (LPRS) analysis showing the $\overline{S}_{\text{LPRS}}$-matrix perturbations from (*j*)-sensors residues *vs.* (*i*)-effector residues based on the previously defined strength of perturbations as indicated by the scales on the right of the maps. The subfigures in the central column represent LPRS analysis results for: **D)** hVDAC1 without ligand, **E)** ATP/hVDAC1 complex, **F)** z-z-SWCNT/hVDAC1 complex considering the best binding-conformation for both ligands in hVDAC1, **G)** hVDAC1 without ligand, **H)** ATP/hVDAC1 complex, and **I)** **z-z-**SWCNT/hVDAC1 complex. The regions blue to orange correspond to weak to moderate local perturbations whereas the regions orange to dark red to strong ones in the ATP-entry-point inter-residue communication (*i*, *j*). Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

The behavior of the local perturbation response of (*i*)-hVDAC1 effector residues provides relevant information regarding the effective communication distances ($^{eff}S_{(ij)}$) for intra-segment Cα-Cα atomic distances, which can be categorized as: *i)* short segments (10–22 residues), *ii)* medium segments (23–26 residues), and *iii)* long segments (27–50 residues) in the N-terminal-(α)-helix segment.[55] Considering the results of local perturbation response of (*i*)-hVDAC1 effector residues in the presence and absence of ligands (z-z-SWCNT and ATP), it appears that z-z-SWCNT affects the inter-residue communication of the ATP-entry-point residues following a long distance-perturbations response, affecting allosteric-phosphorylation residues like SER 104 and TYR 250 or other as HIS 276 and GLY 281. Even though the ATP-ligand has a similar long-distance perturbation response, but the perturbed-residues were different (like GLY 48, TRP 67, ASN 186 and TYR 228). Also, the interaction pattern of z-z-SWCNT is significantly different regarding the physiological condition from the unoccupied ATP-entry-point residues, and the effector-residue like ARG 2 keeps its flexibility properties.

7.3.3 Molecular dynamics results. Presently, computational approaches to model channel nanotoxicity have paramount importance in mitotarget drug-discovery and rational-design of carbon nanomaterials (SWCNT).[12-13] Previous theoretical and *in vitro* studies have demonstrated that SWCNT can induce mitochondrial toxicity with special selectivity by the mitochondrial channels.[56-57] In this context, computational methods like MD can efficiently model with good accuracy many protein-ligand complexes based on the prediction of binding conformations and affinities (FEB values) between drug molecules (*i.e.*, the native hVDAC1-substrate ATP and the z-z-SWCNT) and the mitotarget (*i.e.*, ATP-entry-point). Herein, it is important to note that the *in silico* nanotoxicological assays correspond to a very early stage of the nano-risks evaluation. Therefore, several approximations are adopted in the new field of "Computational Nanotoxicology" during the simulations (modeling z-z-SWCNT/hVDAC1-channel nanotoxicity). In this context, we assume that the probability of occurrence of potential off-target z-z-SWCNT/hVDAC1-interactions with lipid bilayers of cell membranes (including the lipid bilayers of mitochondria membranes as outer-and-inner mitochondrial membranes) in the biological environment could be considered from medium to low. As mentioned in the **Introduction** section the hVDAC1 channel is the most densely localized protein in the outer mitochondrial membrane of all human

cells,[1, 8-10] and on the other hand, the SWCNT (z-z-SWCNT) present high ability to accumulate in the mitochondrial matrix (mitotropic behavior). According to this, several steps biochemical could bring on potential off-target hVDAC1-interactions of z-z-SWCNT with the lipid bilayers of mitochondria membranes like: i) z-z-SWCNT first pass across the outer mitochondrial membrane, where the hVDAC1 channel is embedded. ii) Then, z-z-SWCNT passes to the mitochondrial inter-membrane space following by iii) the passage of z-z-SWCNT to the mitochondrial matrix across the mitochondrial inner-membrane where the ATP-entry-point of hVDAC1 is placed in the facing-inward conformation in mitochondrial matrix.

Then, we show the initial and final MD configurations for the isolated ligand-hVDAC1 systems, where the water molecules and ions were excluded for a clearer view of ligands binding position in the hVDAC1 channel. In the case of ATP-hVDAC1 complex, the ATP molecule does not move much inside the hVDAC1 channel from its docking position, rather prefers to be close to the ATP-entry-point. On the other hand, for the z-z-SWCNT-hVDAC1 complex, we show that the z-z-SWCNT moves toward the hVDAC1 from its docking position during simulation and interacts strongly with the α-helix segment. The MD simulation results are depicted in **Figure 8**.

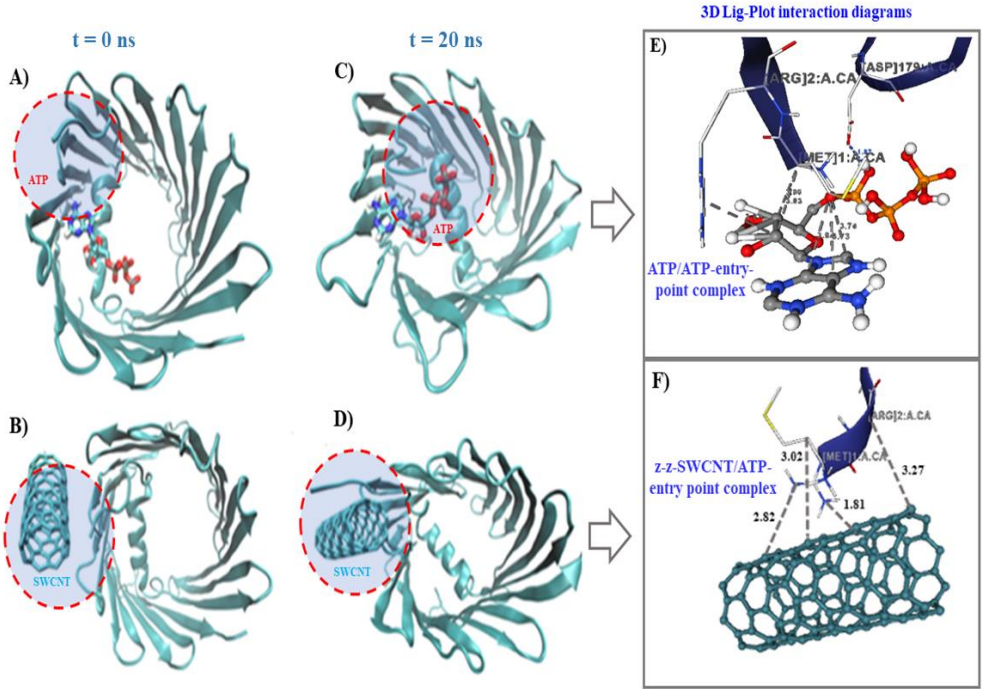


Figure 8. Snapshots of ligand-hVDAC1 complexes obtained from MD simulations (red dotted line). Herein, **A)** and **B)** represent snapshots of the ATP/hVDAC1 complex (marked by red circle) and the z-z-SWCNT/hVDAC1 (marked by red circle),

respectively, at the beginning of the MD simulation (t = 0 ns), whereas **C)** and **D)** represent snapshots from the same complexes at the end of MD simulations (t = 20 ns). The diagrams, **E)** and **F)** represent the 3D-lig-plot hydrophobic interaction with the corresponding critical interatomic distance values ($d_{ij} \leq 7$ Å) from relevant residues of the ATP-entry-point depicted for the aforementioned docking complexes just considering the end of MD simulations (t = 20 ns). Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

To trace the ligand movement from its docked position in the complexes during the simulations, the distances between the center of masses (CM) of the ligand (z-z-SWCNT or ATP) and the ATP-entry-point were calculated and are depicted in **Figure 9.**

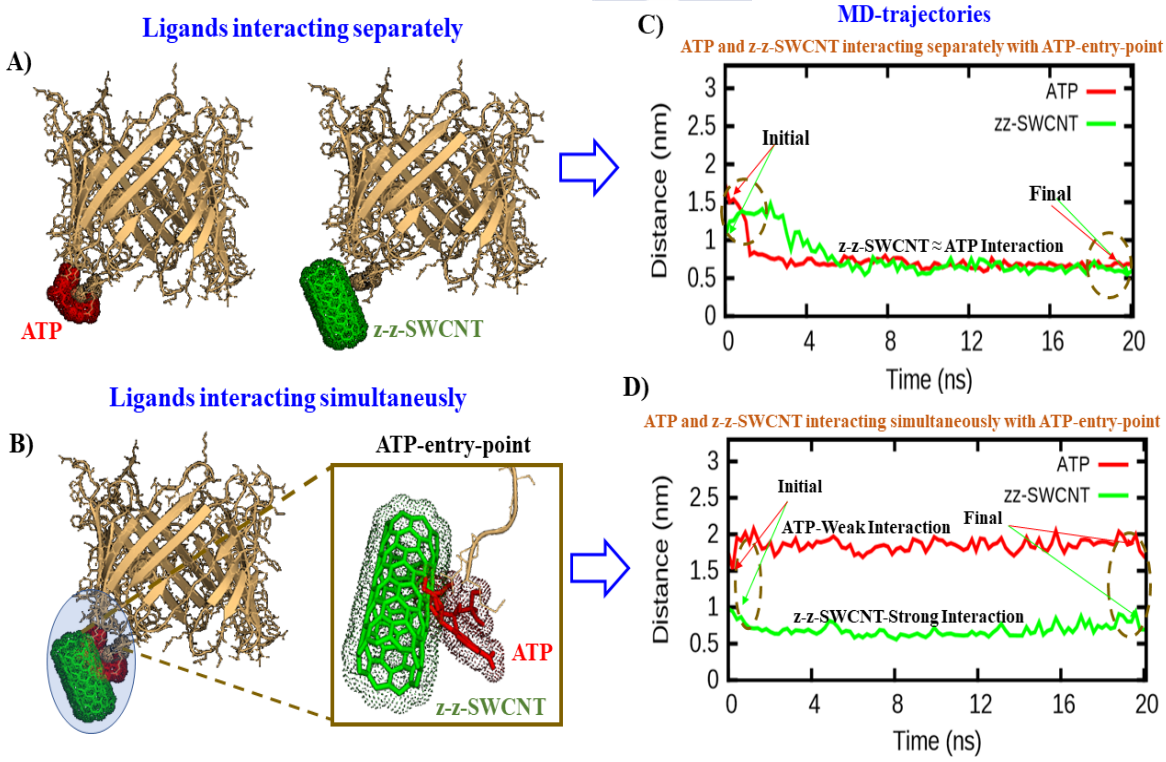


Figure 9. In the left column, 3D-molecular structures are represented for the simulated complexes, namely: **A)** ATP/hVDAC1 and z-z-SWCNT/hVDAC1 complexes, and **B)** Representation of the co-interacting system formed by the ATP molecule and the z-z-SWCNT ligand interacting in the same biophysical environment (ATP-entry point). In the right column, the calculated distances between the center of mass (CM) of the ligands and the CM of ATP-entry-point residues during 20 ns of MD simulations are presented. **C)** MD-trajectories-based distances of the ATP and z-z-SWCNT interacting

separately with the ATP-entry-point, and **D)** MD-trajectories-based distances of ATP and z-z-SWCNT interacting simultaneously with the ATP-entry-point. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

This strategy allowed us to reveal important toxicological aspects based on the selectivity and/or affinity of z-z-SWCNT with mitochondrial channels (hVDAC1).[12-13, 58] According to the obtained results (see **Figure 9C**), the distance between the CM of ATP and the CM of ATP-entry-point in the ATP/ATP-entry-point system decreases from 1.7 nm to 0.65 nm during 20 ns simulation. Similarly, the distance between the CM of z-z-SWCNT and ATP-entry-point in the z-z-SWCNT/hVDAC1 system also decreases from its docked position suggesting a strong attractive interaction between the ligand and the ATP/ATP-entry-point. On the other hand, the distance between the CM of the ligands (ATP and z-z-SWCNT) and the CM of ATP-entry-point (see **Figure 9D**) in ATP + z-z-SWCNT/hVDAC1 system remains almost absolute for both cases. But the distance between the CM of z-z-SWCNT and ATP-entry-point is much shorter than the distance between the CM of ATP and ATP-entry-point, thus implying a stronger attractive interaction between z-z-SWCNT and the ATP-entry-point. This thus suggests that, the z-z-SWCNT-toxicodynamic behavior (*i.e.*, proposed ligand interaction mechanism with the ATP-entry-point) could interfere with the ATP physiological function and in turn, directly affect the mitochondrial bioenergetic function in early stages of ATP-transport through hVDAC1 channel.[3-6]

To have a deeper understanding on the interaction of ligand/ATP-entry-point interactions, the respective free energy of binding (FEB) values were calculated using the Molecular Mechanics Poisson Boltzmann Surface Area algorithm implemented in GROMACS 5.1.4.[49] The obtained total Gibbs free energy of binding (FEB or E_{bind}) of ATP forming the ATP/hVDAC1 complex (–336.82 kcal/mol) presents more negative FEB values, indicating a higher interaction affinity than ATP in the presence of z-z-SWCNT forming the complex of z-z-SWCNT+ATP/hVDAC1 with a FEB = –277.24 kcal/mol (see **Figure 10**). These results suggest a weakening of the ATP affinity when the z-z-SWCNT molecule simultaneously interacts in the same biophysical environment and considering that the obtained free energy of binding of the z-z-SWCNT molecule interacting separately with the ATP-entry-point of hVDAC1 was 70.55 kcal/mol

(**Figure S1**).[12-13, 56-57] From a toxicological point of view, this fact has great significance because the MD results fit with the previous docking results and allow explaining with better clarity the binding interactions by considering the evolution of the affinity in terms of time.

Based on the individual contributions of the different components of FEB, we observed that the molecular mechanics energy (ΔE_{MM}) and non-polar solvation energy (ΔE_{apol}) increase the ATP affinity based on the negative contribution to the obtained FEB values while the polar solvation energy (ΔE_{pol}) weakens the ATP binding affinity due to its positive contribution to FEB. In terms of contributions, ΔE_{MM} potential provides a much larger contribution compared to the polar solvation energy in all the complexes, and the contribution from the non-polar solvation energy is negligibly small. Overall, the free binding energy analysis points out that the ATP molecule has a strong affinity with ATP-entry-point, which weakens in the presence of z-z-SWCNT suggesting potential channel nanotoxicity.[12-13, 56-57] (see **Figure 10**).

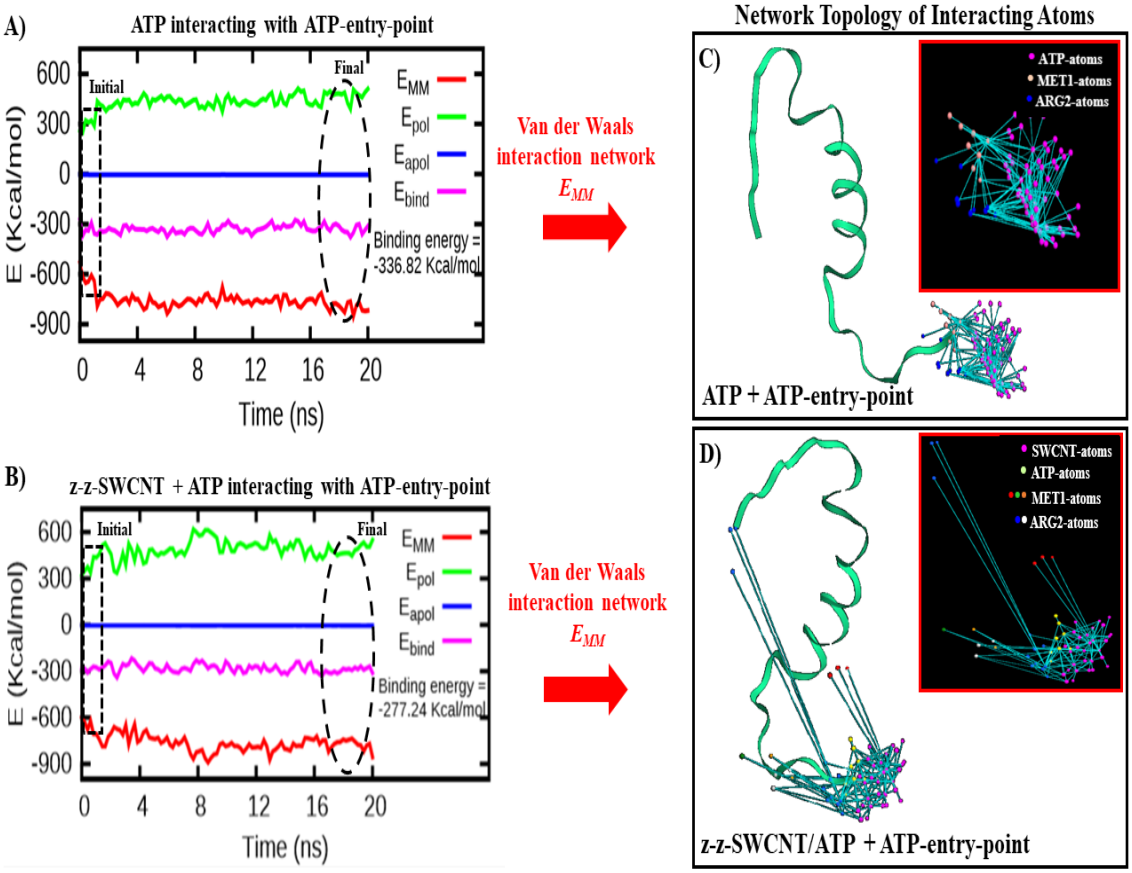


Figure 10. Representation of free binding energy (E_{bind}) decomposition for the complexes: **A)** ATP/hVDAC1 and **B)** co-interacting system as z-z-SWCNT+ATP/hVDAC1 *vs.* simulation time (ns), highlighting the start and end of the

MD simulations. The terms E_{MM}, $E_{pol} \approx G_{pol}$, $E_{apol} \approx G_{apol}$ are the average molecular mechanics potential energy (red line), polar solvation energy (green line), non-polar solvation energy (blue lines), and Gibbs binding free energy (pink line), respectively. Anisotropic network topology performed for the obtained systems based on the best energy contribution (ΔE_{MM}) to the binding free energy as **C)** ATP + ATP-entry-point and **D)** co-interacting system like z-z-SWCNT+ATP/ATP-entry-point. Additional details can be found in the supplementary material like **Figure S1**. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

The computational modeling based on the network approach for van der Waals interactions between atoms of the ligands and the ATP-entry-point reveals the existence of marked differences in the network topology when the two systems: **1)** ATP/hVDAC1 and **2)** z-z-SWCNT+ATP/ hVDAC1 (**Figure 10**; **C** and **D**) are compared. The latter analysis suggests that z-z-SWCNT near the ATP-entry-point induces conformational rearrangements of residues that can affect the physiological function of hVDAC1 in the ATP translocation.

Then, to deepen our understanding of ligand interactions with residues of the ATP-entry-point, we determined the individual contribution of the hVDAC1-residues to the binding free energy by decomposing the total binding energy. For this purpose, the decomposition energy of the complexes per residue is plotted in **Figure 11**. Inspection of **Figure 11A** suggests that the interaction between ATP and the first five residues (like MET1, ARG2, GLY3, SER4, ALA5) of the hVDAC1 which include the ATP-entry-point activate other residue clusters starting from LEU 13 to ARG18 which are part of the surrounding area of the ATP-transport catalytic point, as mentioned above (**Figure 4E**) and also activate the VAL20 to LYS 23 in the middle of the hVDAC1 channel that contributes for the stabilization of ATP molecule in the ATP-hVDAC1 complex during ATP-efflux.

In this context, the results obtained for the interaction of the ATP + hVDAC1 complex (**Figure 11**; **A** and **D**) show that the ATP molecule is able to activates multiple hVDAC1 residues during its interaction with the ATP-entry-point. Indeed, these residues are involved in the subsequent stages of the ATP translocation. However, the obtained results for the hVDAC1 interaction (ATP-entry-point) in the presence of the z-z-SWCNT molecule for the two evaluated conditions like i) z-z-SWCNT + hVDAC1

(**Figures 11; B** and **E**) and ii) the co-interaction of z-z-SWCNT +ATP/ hVDAC1 (**Figure 11**; **C** and **F**) show a significant decrease in the binding affinity in ≈ - 62 kcal/mol in both formed complexes (z-z-SWCNT/hVDAC1 ≈ z-z-SWCNT-ATP/hVDAC1), compared with the ATP/hVDAC1 complex (**Figure 11**; **A** and **D**) which represents the natural interaction due to the ATP molecule is the native substrate of the hVDAC1 (ATP-entry-point) used like reference control of the simulation.

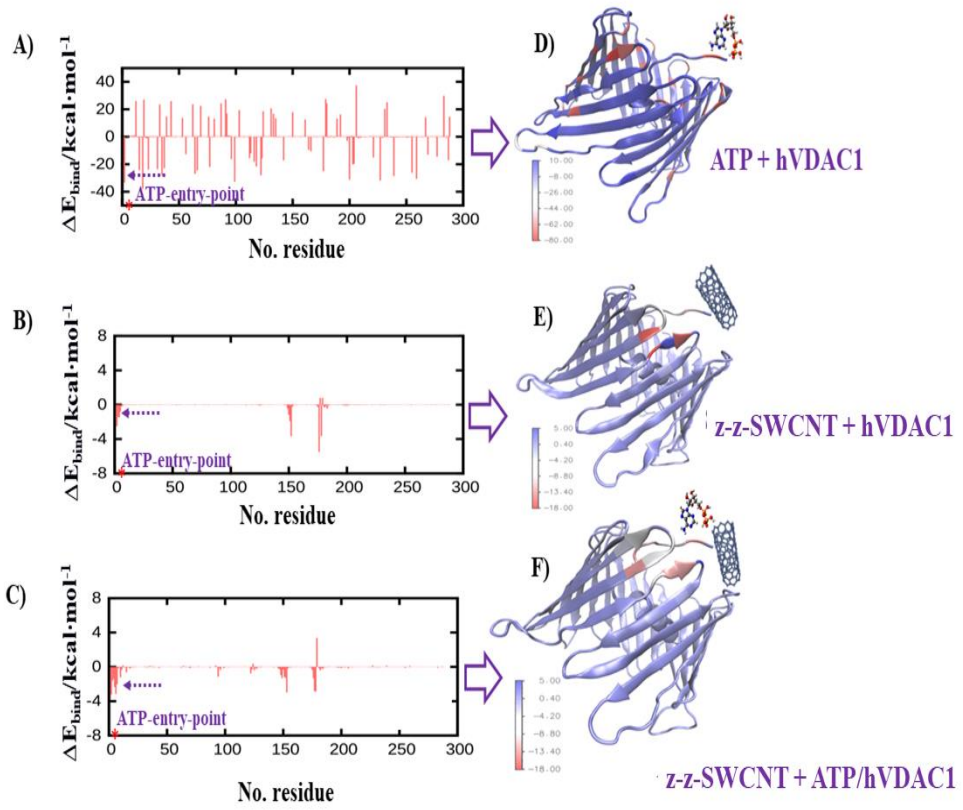


Figure 11. Left column shows a graphical representation of the binding free energy decomposition (ΔE_{bind}) per hVDAC1 residues highlighting the ATP-entry-point residues from the position of MET1 (red asterisk, *) and the corresponding binding free energy ($\Delta E_{bind,}$ like purple arrow dotted lines) for the systems: **A)** ATP/hVDAC1, **B)** z-z-SWCNT/hVDAC1 and **C)** ATP + z-z-SWCNT/hVDAC1. Right column shows a 3D-visualization of the binding energy distribution for complexes like **D)** ATP/hVDAC1, **E)** z-z-SWCNT/hVDAC1, and **F)** ATP+ z-z-SWCNT/hVDAC1. The associated color intensity bar represents ligand interactions (ΔE_{bind}), *i.e.*: weak-interactions (blue), moderate interactions (grey) and strong interactions (red). Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

Next, these results pointing that the isolated and/or simultaneous interaction of the cited ligands (*i.e.*, z-z-SWCNT or co-interaction of z-z-SWCNT + ATP) in the same biophysical environment (ATP-entry-point) can induce drastic changes in the free binding energy patterns (ΔE_{bind}/kcal·mol^{-1}) of the residue clusters involved in the ATP-transport through hVDAC1 channel. For example, decreasing signal transduction of some binding residues to phosphate nucleotides which surrounding the hVDAC1 channel, and range from ASN 242 to SER 244, and the cluster of residues from SER 260 to ASP 267, just to mention some. These results match well with the performed molecular docking results and the ANM approach.

Besides, the proposed approaches fit with the Organization for Economic Co-operation and Development (OECD) guidelines and the International Organization for Standardization on the development of alternative methods in Computational Nanotoxicology.[59,60]

References

1. Bayrhuber, M.; Meins, T.; Habeck, M.; Becker, S.; Giller, K.; Villinger, S.; Vonrhein, C.; Griesinger, C.; Zweckstetter, M.; Zeth, K., Structure of the human voltage-dependent anion channel. *P Natl Acad Sci USA* **2008,** *105* (40), 15370-15375.
2. Guo, X. W.; Mannella, C. A., Conformational Change in the Mitochondrial Channel, Vdac, Detected by Electron Cryomicroscopy. *Biophys J* **1993,** *64* (2), 545-549.
3. Shoshan-Barmatz, V.; Israelson, A.; Brdiczka, D.; Sheu, S. S., The voltage-dependent anion channel (VDAC): Function in intracellular signalling, cell life and cell death. *Curr Pharm Design* **2006,** *12* (18), 2249-2270.
4. Shuvo, S. R.; Ferens, F. G.; Court, D. A., The N-terminus of VDAC: Structure, mutational analysis, and a potential role in regulating barrel shape. *Bba-Biomembranes* **2016,** *1858* (6), 1350-1361.
5. Shoshan-Barmatz, V.; Zakar, M.; Rosenthal, K.; Abu-Hamad, S., Key regions of VDAC1 functioning in apoptosis induction and regulation by hexokinase. *Bba-Bioenergetics* **2009,** *1787* (5), 421-430.
6. Pi, Y. Q.; Goldenthal, M. J.; Marin-Garcia, J., Mitochondrial channelopathies in aging. *J Mol Med* **2007,** *85* (9), 937-951.

7. Zachariae, U.; Schneider, R.; Briones, R.; Gattin, Z.; Demers, J. P.; Giller, K.; Maier, E.; Zweckstetter, M.; Griesinger, C.; Becker, S.; Benz, R.; de Groot, B. L.; Lange, A., beta-Barrel Mobility Underlies Closure of the Voltage-Dependent Anion Channel. *Structure* **2012,** *20* (9), 1540-1549.

8. Ujwal, R.; Cascio, D.; Colletier, J. P.; Faham, S.; Zhang, J.; Toro, L.; Ping, P. P.; Abramson, J., The crystal structure of mouse VDAC1 at 2.3 angstrom resolution reveals mechanistic insights into metabolite gating. *P Natl Acad Sci USA* **2008,** *105* (46), 17742-17747.

9. Choudhary, O. P.; Paz, A.; Adelman, J. L.; Colletier, J. P.; Abramson, J.; Grabe, M., Structure-guided simulations illuminate the mechanism of ATP transport through VDAC1. *Nat Struct Mol Biol* **2014,** *21* (7), 626-632.

10. Okada, S. F.; O'Neal, W. K.; Huang, P. B.; Nicholas, R. A.; Ostrowski, L. E.; Craigen, W. J.; Lazarowski, E. R.; Boucher, R. C., Voltage-dependent anion channel-1 (VDAC-1) contributes to ATP release and cell volume regulation in murine cells. *J Gen Physiol* **2004,** *124* (5), 513-526.

11. Giri, A. K.; Teixeira, F.; Cordeiro, M. N. D. S., Structure and kinetics of water in highly confined conditions: A molecular dynamics simulation study. *J Mol Liq* **2018,** *268*, 625-636.

12. Gonzalez-Durruthy, M.; Werhli, A. V.; Seus, V.; Machado, K. S.; Pazos, A.; Munteanu, C. R.; Gonzalez-Diaz, H.; Monserrat, J. M., Decrypting Strong and Weak Single-Walled Carbon Nanotubes Interactions with Mitochondrial Voltage-Dependent Anion Channels Using Molecular Docking and Perturbation Theory. *Sci Rep-Uk* **2017,** *7*.

13. Gonzalez-Durruthy, M.; Werhli, A. V.; Cornetet, L.; Machado, K. S.; Gonzalez-Diaz, H.; Wasiliesky, W.; Ruas, C. P.; Gelesky, M. A.; Monserrat, J. M., Predicting the binding properties of single walled carbon nanotubes (SWCNT) with an ADP/ATP mitochondrial carrier using molecular docking, chemoinformatics, and nano-QSBR perturbation theory. *Rsc Adv* **2016,** *6* (63), 58680-58693.

14. Wang, X.; Guo, J.; Chen, T.; Nie, H. Y.; Wang, H. F.; Zang, J. J.; Cui, X. X.; Jia, G., Multi-walled carbon nanotubes induce apoptosis via mitochondrial pathway and scavenger receptor. *Toxicol in Vitro* **2012,** *26* (6), 799-806.

15. Noskov, S. Y.; Rostovtseva, T. K.; Chamberlin, A. C.; Teijido, O.; Jiang, W.; Bezrukov, S. M., Current state of theoretical and experimental studies of the voltage-dependent anion channel (VDAC). *Bba-Biomembranes* **2016,** *1858* (7), 1778-1790.

16. Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J., Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc* **2016,** *11* (5), 905-919.

17. Shukla, D.; Hernandez, C. X.; Weber, J. K.; Pande, V. S., Markov State Models Provide Insights into Dynamic Modulation of Protein Function. *Accounts Chem Res* **2015,** *48* (2), 414-422.

18. Briones, R.; Weichbrodt, C.; Paltrinieri, L.; Mey, I.; Villinger, S.; Giller, K.; Lange, A.; Zweckstetter, M.; Griesinger, C.; Becker, S.; Steinem, C.; de Groott, B. L., Voltage Dependence of Conformational Dynamics and Subconducting States of VDAC-1. *Biophys J* **2016,** *111* (6), 1223-1234.

19. Mitternacht, S.; Berezovsky, I. N., Coherent Conformational Degrees of Freedom as a Structural Basis for Allosteric Communication. *Plos Comput Biol* **2011,** *7* (12).

20. Keskin, O.; Durell, S. R.; Bahar, I.; Jernigan, R. L.; Covell, D. G., Relating molecular flexibility to function: A case study of tubulin. *Biophys J* **2002,** *83* (2), 663-680.

21. Greener, J. G.; Sternberg, M. J. E., AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. *Bmc Bioinformatics* **2015,** *16*.

22. Oliwa, T.; Shen, Y., cNMA: a framework of encounter complex-based normal mode analysis to model conformational changes in protein interactions. *Bioinformatics* **2015,** *31* (12), 151-160.

23. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R., Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminformatics* **2012,** *4*.

24. Dresselhaus, G.; Riichiro, S., *Physical properties of carbon nanotubes*. World scientific: 1998.

25. Alfonsi, J.; Meneghetti, M., Small crystal approach for the electronic properties of double-wall carbon nanotubes. *New J Phys* **2009,** *11*.

26. Trott, O.; Olson, A. J., Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem* **2010,** *31* (2), 455-461.

27. Lavecchia, A., Machine-learning approaches in drug discovery: methods and applications. *Drug Discov Today* **2015,** *20* (3), 318-331.

28. Smith, T. C.; Frank, E., Introducing machine learning concepts with WEKA. In *Statistical genomics*, Springer: 2016; pp 353-378.

29. Elokely, K. M.; Doerksen, R. J., Docking Challenge: Protein Sampling and Molecular Docking Performance. *Journal of Chemical Information and Modeling* **2013,** *53* (8), 1934-1945.

30. Sutherland, J. J.; Nandigam, R. K.; Erickson, J. A.; Vieth, M., Lessons in molecular recognition. 2. Assessing and improving cross-docking accuracy. *Journal of Chemical Information and Modeling* **2007,** *47* (6), 2293-2302.

31. da Silveira, C. H.; Pires, D. E. V.; Minardi, R. C.; Ribeiro, C.; Veloso, C. J. M.; Lopes, J. C. D.; Meira, W.; Neshich, G.; Ramos, C. H. I.; Habesch, R.; Santoro, M. M., Protein cutoff scanning: A comparative analysis of cutoff dependent and cutoff free methods for prospecting contacts in proteins. *Proteins* **2009,** *74* (3), 727-743.

32. Liu, K.; Kokubo, H., Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations: A Cross-docking Study. *Journal of Chemical Information and Modeling* **2017,** *57* (10), 2514-2522.

33. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat Rev Drug Discov* **2004,** *3* (11), 935-949.

34. Ballante, F.; Marshall, G. R., An Automated Strategy for Binding-Pose Selection and Docking Assessment in Structure-Based Drug Design. *Journal of Chemical Information and Modeling* **2016,** *56* (1), 54-72.

35. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **2008,** *4* (3), 435-447.

36. Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R., Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput Phys Commun* **1995,** *91* (1-3), 43-56.

37. Daura, X.; Mark, A. E.; van Gunsteren, W. F., Parametrization of aliphatic CHn united atoms of GROMOS96 force field. *J Comput Chem* **1998,** *19* (5), 535-547.

38. van Gunsteren, W. F., *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Biomos ; Zürich: 1996.

39. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The Missing Term in Effective Pair Potentials. *J Phys Chem-Us* **1987,** *91* (24), 6269-6271.

40. Giri, A. K.; Spohr, E., Conformational Equilibria of Organic Adsorbates on Nanostructures in Aqueous Solution: MD Simulations. *J Phys Chem C* **2015,** *119* (45), 25566-25575.

41. Giri, A. K.; Spohr, E., Influence of Chain Length and Branching on the Structure of Functionalized Gold Nanoparticles. *J Phys Chem C* **2018,** *122* (46), 26739-26747.

42. Giri, A. K.; Spohr, E., Cluster formation of NaCl in bulk solutions: Arithmetic vs. geometric combination rules. *J Mol Liq* **2017,** *228*, 63-70.

43. Hockney, R. W.; Eastwood, J. W., *Computer Simulation Using Particles*. CRC Press: 1988.

44. Parrinello, M.; Rahman, A., Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. *J Appl Phys* **1981,** *52* (12), 7182-7190.

45. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* **2007,** *126* (1).

46. Giri, A. K.; Teixeira, F.; Cordeiro, M. N. D., Salt separation from water using graphene oxide nanochannels: A molecular dynamics simulation study. *Desalination* **2019,** *460*, 1-14.

47. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J Chem Phys* **1995,** *103* (19), 8577-8593.

48. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **1993,** *98* (12), 10089-10092.

49. Kumari, R.; Kumar, R.; Consortium, O. S. D. D.; Lynn, A., g_mmpbsa - A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of chemical information and modeling* **2014,** *54* (7), 1951-1962.

50. Jimenez, J.; Doerr, S.; Martinez-Rosell, G.; Rose, A. S.; De Fabritiis, G., DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics* **2017,** *33* (19), 3036-3042.

51. Feinstein, W. P.; Brylinski, M., Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. *J Cheminformatics* **2015,** *7*.

52. Meslamani, J.; Rognan, D.; Kellenberger, E., sc-PDB: a database for identifying variations and multiplicity of 'druggable' binding sites in proteins. *Bioinformatics* **2011,** *27* (9), 1324-1326.

53. Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **2010,** *66*, 12-21.

54. Chennubhotla, C.; Bahar, I., Signal propagation in proteins and relation to equilibrium fluctuations. *Plos Comput Biol* **2007,** *3* (9), 1716-1726.

55. Ikeda, K.; Hirokawa, T.; Higo, J.; Tomii, K., Protein-segment universe exhibiting transitions at intermediate segment length in conformational subspaces. *Bmc Struct Biol* **2008,** *8*.

56. Park, K. H.; Chhowalla, M.; Iqbal, Z.; Sesti, F., Single-walled carbon nanotubes are a new class of ion channel blockers. *J Biol Chem* **2003,** *278* (50), 50212-50216.

57. Chhowalla, M.; Unalan, H. E.; Wang, Y. B.; Iqbal, Z.; Park, K.; Sesti, F., Irreversible blocking of ion channels using functionalized single-walled carbon nanotubes. *Nanotechnology* **2005,** *16* (12), 2982-2986.

58. Vogele, M.; Kofinger, J.; Hummer, G., Molecular dynamics simulations of carbon nanotube porins in lipid bilayers. *Faraday Discuss* **2018,** *209*, 341-358.

59. OECD Principles for the validation, for regulatory porposes of (Quantitative) Structure Activity Relationship Model, http://www.oecd.org/, accessed **13/03/2020**.

60. ISO/TC 229, *Nanotechnology*, **2011**, 1–11, Draft 4.

Supplementary material

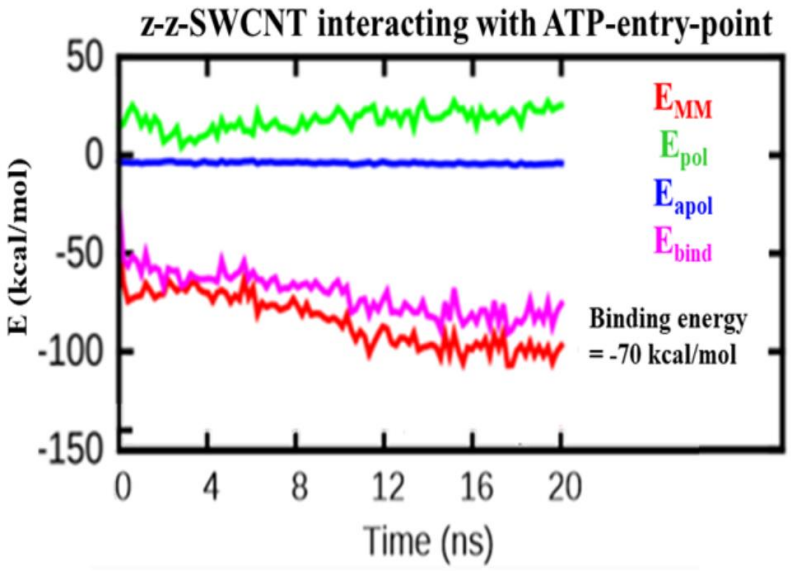


Figure S1. Representation of free binding energy (E_{bind}) decomposition of the z-z-SWCNT interacting individually with the ATP-entry-point of the hVDAC1 channel as reference control of simulation for comparison purposes. Source Elsevier article, please visit: https://doi.org/10.1016/j.nantod.2020.100913 (accessed on october 2020)

Chapter 8.

Michael González-Durruthy, Riccardo Concu, Juan M. Ruso, and M. Natália D.S. Cordeiro. New Mechanistic Insights on Carbon Nanotubes' Nanotoxicity Using Isolated Submitochondrial Particles, Molecular Docking, and Nano-QSTR Approaches. Biology, 10, 171, 2021. doi.org/10.3390/biology10030171.

Journal Impact Factor (JIF) (2021): 5.079.

CiteScore (2021): 3.48

Author Contribution M.G.-D.: Conceptualization, methodology-based on structure-based docking virtual screening coupled QSTR/QSAR computational models, performing in vitro experimental validations by using isolated rat-liver submitochondrial particles (RL-SMP), performing carbon nanotubes characterization using Transmission Electron Microscope (TEM) images, performing experimental determination of mitochondrial F0F1-ATPase inhibition in RL-SMP, performing LPRS maps coupled with fractal analysis using box-counting approaches, and writing—original draft preparation. *Corresponding author

Journal authorization:

Chapter 8.

New Mechanistic Insights on Carbon Nanotubes' Nanotoxicity Using Isolated Submitochondrial Particles, Molecular Docking, and Nano-QSTR Approaches.

Abstract: Single-walled carbon nanotubes can induce mitochondrial F0F1-ATPase nanotoxicity through inhibition. To completely characterize the mechanistic effect triggering the toxicity, we have developed a new approach based on the combination of experimental and computational study, since the use of only one or few techniques may not fully describe the phenomena. To this end, the in vitro inhibition responses in submitochondrial particles (SMP) was combined with docking, elastic network models, fractal surface analysis, and Nano-QSTR models. In vitro studies suggest that inhibition responses in SMP of F0F1-ATPase enzyme were strongly dependent on the concentration assay (from 3 to 5 μg/mL) for both pristine and COOH single-walled carbon nanotubes types (SWCNT). Besides, both SWCNTs show an interaction inhibition pattern mimicking the oligomycin A (the specific mitochondria F0F1-ATPase inhibitor blocking the c-ring F0 subunit). Performed docking studies denote the best crystallography binding pose obtained for the docking complexes based on the free energy of binding (FEB) fit well with the in vitro evidence from the thermodynamics point of view, following an affinity order such as: FEB (oligomycin A/F0-ATPase complex) = −9.8 kcal/mol > FEB (SWCNT-COOH/F0-ATPase complex) = −6.8 kcal/mol ~ FEB (SWCNT-pristine complex) = −5.9 kcal/mol, with predominance of van der Waals hydrophobic nano-interactions with key F0-ATPase binding site residues (Phe 55 and Phe 64). Elastic network models and fractal surface analysis were performed to study conformational perturbations induced by SWCNT. Our results suggest that interaction may be triggering abnormal allosteric responses and signals propagation in the inter-residue network, which could affect the substrate recognition ligand geometrical specificity of the F0F1-ATPase enzyme in order (SWCNT-pristine > SWCNT-COOH). In addition, Nano-QSTR models have been developed to predict toxicity induced by both SWCNTs, using results of in vitro and docking studies. Results show that this method may be used for the fast prediction of the nanotoxicity induced by SWCNT, avoiding time- and money-consuming techniques. Overall, the

obtained results may open new avenues toward to the better understanding and prediction of new nanotoxicity mechanisms, rational drug design-based nanotechnology, and potential biomedical application in precision nanomedicine.

8.1 Introduction

The coupled mechanical co-rotating between the γ and ε subunits that form the mitochondrial F1-ATP synthase (complex V) favors the H^+ protons flux necessary for ATP synthesis in all eukaryotic cells [1,2]. This bioenergetic process involves several synchronized conformational changes which are critical for the survival or death of the cells [1]. In this regard, a few years ago, it was shown that under pathological conditions like chronic diseases such as cancer, Alzheimer's disease, Parkinson's disease, and mitochondrial encephalopathy, lactic acidosis (MELAS) syndrome, several toxic events, including nanotoxicity induced by single walled carbon nanotubes (SWCNT), may trigger F0F1ATPase dysfunction [3,4]. As a consequence, the ATP cellular reserves are abruptly consumed by a reverse biochemical reaction which paradoxically hydrolyses significant amounts of ATP, compromising the cellular homeostasis and viability [3,5,6]. Several chemical agents (including carbon nanoparticles) have shown a high affinity/selectivity by the bioenergetic mechanisms based on ATP hydrolysis, particularly nanoparticle-based single-walled carbon nanotubes (SWCNTs), which have been studied by their selective nanotoxicity effects on mitochondria (mitotropic behavior) [7–10].

To the best of our knowledge, the toxicological modulation of mitochondrial ATP bioenergetic mechanisms released by the exposure with SWCNT-pristine and oxidized-SWCNT (SWCNT-COOH) have been insufficiently characterized in order to explain the mitochondrial nanotoxicity induced by SWCNT. On the other hand, this mechanistic knowledge could be very useful to implement strategies on the named "precision mitochondrial nanomedicine" to improve selectivity for the treatment of brain, cardiac diseases, and cancer using the mitotropic behavior of SWCNT to address active pharmacological principles as new targeting of the mitochondrial F0F1-ATPase [9–14]. In this context, we hypothesize that SWCNT-pristine could act by mimicking the pharmacodynamic behavior of the Oligomycin A, which is the specific inhibitor of the mitochondrial ATP-hydrolysis that modulates the activity of the c-ring-F0-ATP hydrolase subunit. However, in the case of the SWCNT-COOH, the F0-ATPase binding interaction could be more attenuated by the presence of the carboxyl group.

From the structural point of view, the c-ring-F0-ATP hydrolase subunit represents an uncoupling channel which is part of the mitochondrial permeability transition pore-induced association to mitochondrial dysfunction and apoptosis [15,16]. Following this idea, we suggest that SWCNT could promote the selective inhibition of the F0-ATPase under pathological conditions like cancer where the F0F1-ATPase activity is abnormally exacerbated [16–19].

In this regard, computational approaches like molecular docking simulation, elastic network models, fractal surface approaches linked to nano-quantitative structure–activity/toxicity relationships (Nano-QSAR/QSTR models), and others [9,20–23], could be efficiently applied to the exhaustive exploration of the underlying mechanisms of mitochondrial bioenergetic dysfunction (pathological ATP-hydrolysis) from the structural point of view for therapeutic purposes.

Protein structures cannot be investigated using the classical Euclidian mathematical approach. Due to this nature, surface and protein's chain should be studied using the fractal approach. It is well-known that the fractal dimensions (FDs) are directly associated to the backbone non-Euclidean geometry, as well as to the irregular geometric nature and fractal surface properties of the binding sites (ATPaseF0F1 binding sites). This is explained by the fact that most of the ligand–protein binding processes occur under strict conditions of specificity and, at the same time, that these thermodynamic processes depend on surface phenomena with a defined geometric pattern of stereospecificity and complementarity with the cited binding sites [24]. For this instance, we thus suggest that small changes in the fractal geometry-based surface patterns could directly affect not only the native ATPaseF0F1 binding sites' folding and solvent accessible surface in the unbound state (unoccupied ATPaseF0F1), but also the conformational entropy and thermodynamic stability of the formed docking complexes generated between the ATPaseF0F1 and the different single-walled carbon nanotubes tested [24,25]. In addition, elastic network models may also be used to study proteins since they may be able to predict global dynamics of proteins and proteins' complexes [26–28]. Thus, these two methods can be used, together with other methods, to study conformational changes induced by SWCNT that may produce harmful effects, inactivation, and so on.

Another relevant approach to study toxicity is computational nano-quantitative structure–activity/toxicity relationships (nano-QSAR/QSTR), which are essential tools to support the discovery process of toxicological effects of nanomaterials (SWCNT).

Several approaches have been developed and applied recently to predict potential harmfulness of nanoparticles and nanomaterials [9–14,29]. These in silico tools have the quality of being versatile and reconfigurable to many problems. For example, the nano-quantitative structure–binding relationship (Nano-QSBR) models are a type of Nano-QSTR which are able to associate the physico-chemical properties of nanomaterials (nano-descriptors) with the theoretical free energy of binding (FEB values, kcal/mol) obtained from the molecular docking studies and also to experimental nanotoxicological outputs [13,14,30].

Due to this, the QSAR (Nano-QSTR) paradigm has been applied since the beginning of the "nano revolution" as a useful methodology able to support toxicity profiling of nanomaterials and CNT [31–35]. Several approaches by many authors have been reported combining different molecular descriptors, methodologies, and algorithms, including machine learning and deep learning [34–42]. In this sense, it is strongly advisable to use Nano-QSTR approaches while performing toxicity profiling of CNT and nanomaterials, since they may be able to predict toxicity as well as directly correlate toxicity/activity with specific features of nanomaterials. In addition, in silico approaches are strongly encouraged by national and supranational authorities in the light of the European Union (EU) 3R principles (replacement, reduction, refinement). Currently, the main limitation of these computational methods is to address a feasible mechanistic interpretation of the nanotoxicity phenomena at the atomic level, in many cases [43].

In this work, we propose for the first time a combination of computational modeling approaches, based on molecular docking simulations, elastic network models, fractal surface approaches, and Nano-QSTR calculations, along with experimental validation to tackle the study of binding interactions between single-walled carbon nanotubes with the mitochondrial F0F1-ATPase to contribute to the rational drug design-based nanotechnology, mitotarget drug discovery, and the new area of precision mitochondrial nanomedicine.

8.2 Materials and Methods

8.2.1. Experimental Section

8.2.1.1 Reagents and Solutions

Sucrose, ethylene-glycol-bis (b-aminoethyl)-N,N,N0,N0-tetraacetic acid (EGTA), potassium succinate (plus 2 mM rotenone), K2HPO4, and piperazine-N-2-ethanesulfonic acid (Hepes), dimetilsulfóxido (DMSO), and Biuret reagent. All other reagents were

commercial products of the highest purity grade available. Single-walled carbon nanotubes like SWCNT-pristine and carboxylated-CNT (SWCNT-COOH) with very low conductivity and semi-metallic properties were provided by Cheaptubes Company (http://cheaptubes.com/shortohcnts.htm) for the execution of experimental in vitro assays using submitochondrial particles. All other reagents were commercial products of the highest purity grade available and were purchased from Sigma-Aldrich products

8.2.1.2 Carbon Nanotubes' Characterization

For this instance, a Transmission Electron Microscope (TEM, Tecnai G2-12-SpiritBiotwin FEI-120 kV) was used to characterize the morphology of SWCNT-pristine and oxidized carbon nanotubes such as SWCNT-COOH. The CNT were synthesized by using a catalytic chemical vapor deposition (CCVD) method and functionalized using a concentrated acid mixture of H2SO4:HNO3 mixed (2:1). On the other hand, in order to discover the molecular mechanisms of interaction inhibition of the carbon nanotubes with the F0-ATPase, two types of single-walled carbon nanotubes (SWCNT-pristine and SWCNT-COOH) were modeled by using the Avogadro software, which can be efficiently applied as an advanced molecule editor and visualizer for molecular modeling and computational chemistry. Herein, it is important to note that the in silico analysis was performed just for the purpose of proposing a theoretically rigorous mechanism to explain the potential inhibition of the single-walled carbon nanotubes used on the F0-ATP-ase inhibition. For this reason, the theoretically modeled SWCNTs should not be taken as exact copies from the structural point of view compared with the experimentally tested CNT (SWCNT-pristine and SWCNT-COOH) used in in vitro assays. In this sense, for computational purposes, several approximations were performed mainly based on the diameter and length of carbon nanotubes theoretically modeled compared with those experimentally evaluated, see **Figure 1**.

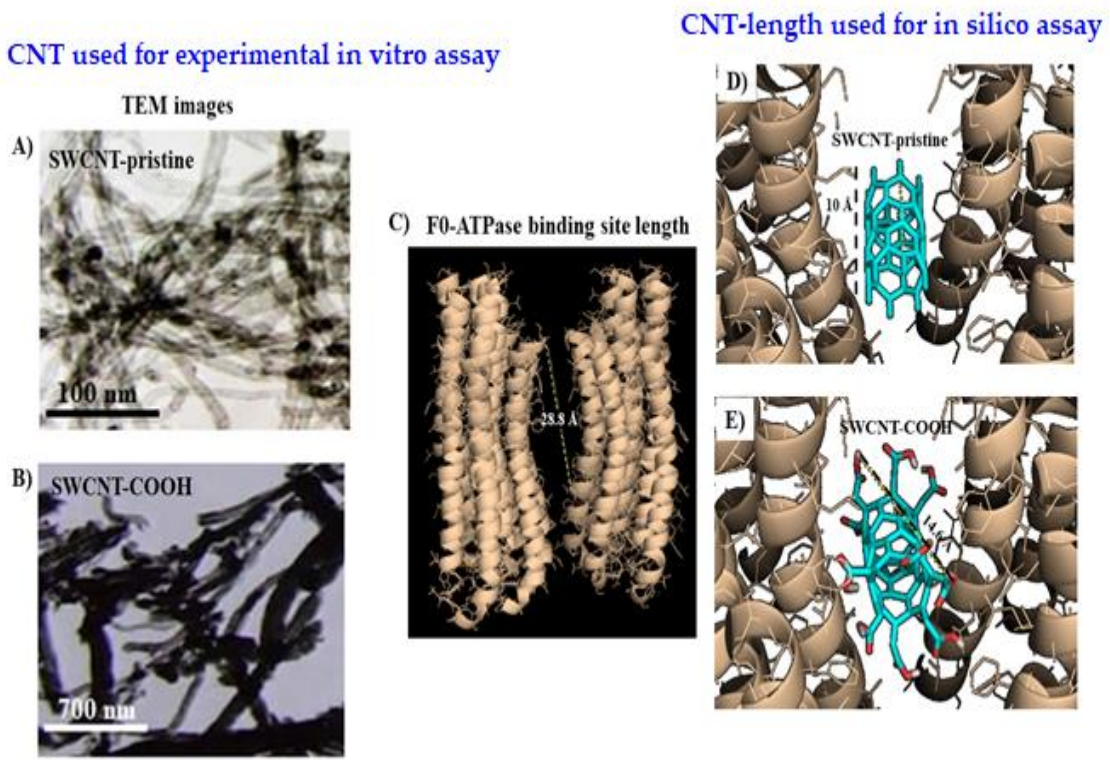


Figure 1. On the right, Transmission Electron Microscope (TEM) images obtained of carbon nanotubes, such as (A) SWCNT-pristine and (B) SWCNT-COOH used in this study for the experimental in vitro assay. On the left, (C) representation of the length of the unoccupied F0-ATPase binding site, (D) and (E) representation of the lengths of the theoretically modeled SWCNT-pristine and SWCNT-COOH within the F0-ATPase used for the in silico assay of F0-ATPase inhibition. Additional details can be found in the Supplementary Information in **Figure S1**. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

8.2.1.3 Isolation of Rat Liver Submitochondrial Particles (SMP)

The frozen rat liver mitochondria (RLM) pellet was thawed and diluted with homogenization medium to contain 20 mg of protein/mL. The mitochondrial suspension was subjected to sonic oscillation four times for 15 s with 30 s intervals, using 80 watts at 4 °C [44–47]. The suspension was then centrifuged at 9750 × g for 10 min at 4 °C and the submitochondrial particles in the supernatant were isolated by additional centrifugation in a Sorval SV-80 vertical rotor for one hour at 15,000 rev/min at 4 °C, using discontinuous gradient containing 1 mL of 0.5 M sucrose and 1 mL of 2.0 M sucrose in 5 mM Tris-HCl, pH 7.4. Finally, the SMP were suspended in the isolation medium, and the final volume was adjusted to give a stock suspension containing 1 mg of protein/mL.

8.2.1.4 Standard Incubation Procedure

Mitochondria liver was isolated and submitochondrial particles (SMP) were energized with 5 mM of potassium succinate (plus 2.5 µM of rotenone) in a standard incubation medium consisting of 125 mM of sucrose, 65 of mM KCl, 2 mM of inorganic phosphate (K_2HPO_4), and 10 mM of HEPES-KOH, pH 7.4, at 30 °C [44–47].

8.2.1.5 Determination of Mitochondrial F0F1-ATPase Inhibition in Isolated Rat Liver Submitochondrial Particles (SMP)

Isolated rat liver submitochondrial particles (isolated-F0F1-ATPase) (20 mg of protein) were incubated according to the following experimental groups: (1) untreated SMP, (2) SMP + DMSO (100 mM), (3) SMP + SWCNT samples (SWCNT-pristine, SWCNT-COOH) in the range of concentration of 0.5–5 µg/mL, (4) SMP + Oligomycin A (1 µM) as a positive control, and (5) SMP + Oligomycin A (1 µM) + SWCNT samples at 5 µg/mL as an additional control assay. The reactions are started by addition of enzyme, such as H^+-c-ring/F_0-ATPase (80 µg of protein). The total volume was 1 mL. After 10 min at 37 °C, the reaction was stopped by addition of 0.5 trichloroacetic acid, 30% (w/v). Phosphate released by ATP hydrolysis is measured on 0.5 mL of molybdate reagent (10 mM ammonium molybdate in 2.5 M sulfuric acid), 1 mL of acetone, and 0.5 mL of 0.4 M citric acid. After each addition, the tubes are homogenized for 10 s in a vortex mixer. The mitochondrial F0F1-ATPase inhibition (F0-ATPase inhibition) for each treatment was calculated by measuring the absorbance at 355 nm [44–47]. Before all spectrophotometric F0-ATPase inhibition measurements, the blanks with each SWCNT were run and interference absorbance peaks of SWCNT were not observed at 300–400 nm [44–47]. Furthermore, each SWCNT sample was added under continuous stirring by using magnetic stirrer cuvettes with the aim of preventing the agglomeration process for the SWCNTs during the F0F1-ATPase inhibition assay. For this instance, a tip-sonication regime during 5–10 min was applied which prevents the SWCNT exfoliation into individual SWCNT samples (SWCNT-pristine, SWCNT-COOH), generating a non-agglomerated suspension in monodisperse state before exposure to submitochondrial particle suspension [48–50].

8.2.1.6 Statistical Procedures for the Mitochondrial Assays Using SMP

The one-way analysis of variance (ANOVA) followed by a post hoc Newman–Keuls multiple comparison test was used in order to determine statistical differences between F0-ATPase inhibition assays as independent unrelated experimental groups. In this context, the Newman–Keuls test was used as a multiple and tiered comparison procedure to identify experimental group statistical means that are significantly different from each other from the different experimental conditions evaluated, namely: (i) untreated submitochondrial particles control (SMP as F0-ATPase), (ii) DMSO-treated SMP, (iii) CNT-treated SMP (i.e., SWCNT-pristine or SWCNT-COOH at 1–5 μg/mL), (iv) Oligomycin A-treated SMP (Oligomycin A is a specific F0F1-ATPase inhibitor used as a positive control), and (v) treated SMP mixed with SWCNT or SWCNT-COOH at concentration of 5 μg/mL + Oligomycin A (1 μM). All the biochemical tests, by using isolated rat liver mitochondria (RLM) and submitochondrial particles (SMP), were performed at least three times in triplicate. Normality and variance homogeneity were verified using Shapiro–Wilks and Levene tests respectively, before using one-way ANOVA. In all cases, significance level was set at 5%.

8.2.2. Theoretical Section

8.2.2.1. Molecular Docking Study

Docking simulations were performed using Autodock tools mixed Autodock Vina to understand the strength of biochemical interactions across CNT family members (SWCNT-pristine and oxidized-CNT (SWCNT-COOH)) and oligomycin A on F0-ATPase. These in silico binding interactions were performed only to explain hidden biophysical and pharmacodynamic mechanisms observed in the mitochondrial in vitro assays. For this instance, only two types of single walled zigzag SWCNTs (Hamada index $n = 8$, $m = 0$) were modeled, like SWCNT (8.0) and SWCNT-COOH (8.0) as F0F1-ATPase ligands in order to reproduce and model some critical experimental conditions from CNT-properties, like CNT-functionalization linked to observed F0-ATPase inhibition (ATP-hydrolysis inhibition) in isolated RLM and isolated SMP. Following this idea, the F0F1-ATPase C10 ring with oligomycin A from yeast (*Saccharomyces cerevisiae*) as the receptor (protein data bank (PDB) ID: 5BPS, Resolution 2.1Å) was obtained from the *RCSB* Protein Data Bank (PDB) [51]. It is important to note that c-ring-F0-ATPase subunit PDB X-ray structure from *Saccharomyces cerevisiae* (5BPS) can be used in the context of the present docking approaches, taking into account that mitochondrial c-ring-F0-ATPase subunit

PDB X-ray structure from *Rattus norvegicus* with oligomycin A has not been crystallized and included in the *RCSB* PDB [51]. However, the oligomycin A pharmacodynamics mechanism is highly conserved in *Rattus norvegicus* according to previous experimental evidences [17].

Before the molecular docking, ATPase C10 ring molecular structure was optimized using the AutoDock Tools 4 software for AutoDock Vina. The algorithm includes the removal of crystallographic water molecules and all the co-crystallized ATPase C10 ring ligand molecules, such as oligomycin A (Oligo A: $C_{45}H_{74}O_{11}$ ID: EFO) from ATPase C10 ring chains (B, E, K, L, M, O). Oligomycin A is a recognized classical inhibitor of F0F1-ATPase inhibition and it was used as a control to compare the affinity and/or relevant interactions by the re-docking procedure.

This theoretical algorithm was performed to the c-ring F0-ATPase subunit using a grid box size with dimensions of X = 22 Å, Y = 22 Å, and Z = 22 Å, and the c-ring F0-ATPase subunit grid box center X = 19.917 Å, Y = 19.654 Å, and Z = 29.844 Å to evaluate the interaction of SWCNT + c-ring–F0-ATPase [52], considering the oligomycin A environment to evaluate the SWCNT-surface affinity in the c-ring F0-ATPase subunit active binding site.

The docking free energy of binding output results (or FEB values) is defined by affinity (like ΔG_{bind} values) for all docked poses of the formed complexes (SWCNT-F0ATPase) and includes the internal steric forces of a given ligand (SWCNT), which can be expressed as the sum of individual molecular mechanics terms of standard chemical potentials as: van der Waals interactions (ΔG_{vdW}), hydrogen bond ($\Delta G_{H\text{-}bond}$), electrostatic interactions ($\Delta G_{electrost}$), and intramolecular interactions ($\Delta G_{internal}$) ligands (SWCNTs) from empirically validated Autodock Vina scoring function based on default Amber force-field parameters [20–22].

8.2.2.2 Local Perturbation Response Induced by SWCNT on F0-ATPase Subunit

In parallel with docking simulation, a new elastic network model was performed to propose a potential mechanism based on the SWCNT propensity to perturb the intrinsic motion of F0-ATPase subunit binding residues involved in the docking interactions. For this purpose, the F0-ATPase is represented as a network or graph of the inter-residue contacts from Cα-F0-ATPase atoms of a residue and the overall potential is simply the sum of harmonic potentials between interacting nodes (F0-ATPase residues). The network includes all interactions within a cutoff distance < 4 Å. Information about the orientation of each interaction with respect to the global coordinates system is considered within the

force constant matrix and allows prediction of perturbed anisotropic motions [53]. The force constant of the F0-ATPase protein system can be described by a Kirchhoff or Hessian matrix ($H_{i,j}$) to evaluate potential perturbations induced by the SWCNT ligand in the transduction properties of the F0-ATPase enzyme according to the following **equation (1)**:

$$H_{ij} = \begin{bmatrix} H_{1,1} & H_{1,2} \dots\dots\dots\dots\dots\dots\dots\dots\dots & H_{1,N} \\ H_{2,1} & H_{2,2} \dots\dots\dots\dots\dots\dots\dots\dots\dots & H_{2,N} \\ . \quad . & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & . \\ H_{N,1} & H_{N,2} \dots\dots\dots\dots\dots\dots\dots\dots\dots & H_{N.N} \end{bmatrix} \quad (1)$$

where each $H_{i,j}$ is a 3 × 3 matrix which holds the anisotropic information regarding the orientation of residues (i, j nodes). Each such sub-matrix (or the "super element" of the $H_{i,j}$ Hessian matrix) is defined by the **equation (2)** as:

$$H_{ij} = \begin{bmatrix} \delta^2V/\delta X_i\delta X_j & \delta^2V/\delta X_i\delta Y_j & \delta^2V/\delta X_i\delta Z_j \\ \delta^2V/\delta Y_i\delta X_j & \delta^2V/\delta Y_i\delta Y_j & \delta^2V/\delta Y_i\delta Z_j \\ \delta^2V/\delta Z_i\delta X_j & \delta^2V/\delta Z_i\delta Y_j & \delta^2V/\delta Z_i\delta Z_j \end{bmatrix} \quad (2)$$

The second partial derivatives are the harmonic potentials, V, between interacting F0-ATPase residues. These partial derivatives are formed by a simple matrix of cosines and the off-diagonal super elements of the $H_{i,j}$ Hessian matrix are calculated according to **Equation (3)** as:

$$H_{i,j} = \begin{bmatrix} -\gamma\frac{(X_j-X_i)(X_j-X_i)}{S_{i,j}^2} & -\gamma\frac{(X_j-X_i)(y_j-Y_i)}{S_{i,j}^2} & -\gamma\frac{(X_j-X_i)(Z_j-Z_i)}{S_{i,j}^2} \\ -\gamma\frac{(Y_j-Y_i)(X_j-X_i)}{S_{i,j}^2} & -\gamma\frac{(Y_j-Y_i)(Y_j-Y_i)}{S_{i,j}^2} & -\gamma\frac{(Y_j-Y_i)(Z_j-Z_i)}{S_{i,j}^2} \\ -\gamma\frac{(Z_j-Z_i)(X_j-X_i)}{S_{i,j}^2} & -\gamma\frac{(Z_j-Z_i)(Y_j-Y_i)}{S_{i,j}^2} & -\gamma\frac{(Z_j-Z_i)(Z_j-Z_i)}{S_{i,j}^2} \end{bmatrix} \quad (3)$$

where $\gamma = 0.5$ is an interaction constant. The $s_{i,j}$ is the instantaneous distance between nodes or residues *i* and *j*. The diagonal super elements are calculated by the **Equation (4)**:

$$H_{i,i} = -\sum_{j=1, j\neq i}^{N} H_{i,j} \qquad (4)$$

Herein, the force constant matrix $H_{i,j}$ holds information regarding the F0-ATPase-residues position/orientation. The inverse of the Hessian matrix is the covariance matrix of *3N* multi-variant Gaussian distribution, where *p* is an empirical parameter according to the Equation (5) for the new off-diagonal elements of the Hessian matrix which hold the desired information on the residue fluctuations, including the F0-ATPase binding site residues (i, j) involved in the SWCNT-F0-ATPase docking interactions.

$$H_{ij} = -\frac{1}{S_{i,j}^{p+2}} \begin{bmatrix} (X_j - X_i)(X_j - X_i) & (X_j - X_i)(Y_j - Y_i) & (X_j - X_i)(Z_j - Z_i) \\ (Y_j - Y_i)(X_j - X_i) & (Y_j - Y_i)(Y_j - Y_i) & (Y_j - Y_i)(Z_j - Z_i) \\ (Z_j - Z_i)(X_j - X_i) & (Z_j - Z_i)(Y_j - Y_i) & (Z_j - Z_i)(Z_j - Z_i) \end{bmatrix} (5)$$

Then, we tackle the construction of the local perturbation response scanning maps (LPRS maps) by setting the following conditions: (i) unbound F0-ATPase as the control simulation experiment, (ii) oligomycin A + F0-ATPase, (iii) SWCNT-pristine + F0-ATPase, and (iv) SWCNT-COOH + F0-ATPase.

8.2.2.3 Performing Nano-QSTR Approaches

The Nano-QSTR models have been developed using a linear regression approach to predict the mitochondrial F0F1-ATPase inhibition values of the SWCNT studied herein. The values used for the development of the continuous model were obtained from molecular docking experiments considering the free energy of binding (FEB values) obtained from the complexes SWCNT-pristine/F0-ATPase and SWCNT-COOH/F0-ATPase. For this purpose, two different sets for both ligands (SWCNT-pristine, SWCNT-COOH) were efficiently built. Considering the three recognized categories of geometric topologies as: zigzag-SWCNT (Hamada index m = 0, n > 0), amchair-SWCNT (Hamada index m = n), and chiral-SWCNT, characterized by the Hamada index (n, m), with m > 0 and m ≠ n, and with its enantiomers (or mirror images), presenting the Hamada index (m, n), which is different from (n, m), with no reflection symmetry [13,14]. Then, regression Nano-QSTR models were developed using the linear regression tool implemented in the Statistica® suite.

The validation of the Nano-QSTR model was performed using the cross-validation module implemented in the software. This procedure is aimed at assessing the predictive accuracy of a model. The test randomly split the dataset into a training set and a validation

set, ensuring that if an entry was included in the test set it could not be used in the validation set. In so doing, the model was developed using the cases in the training or learning sample, which, in our study, was 70% of the dataset. The predictive accuracy was then assessed using the remaining 30% of the dataset. In addition, we have also reported the applicability domain (AD) for both models.

Finally, the performance of the model was evaluated using the residuals, *R* and R^2, and other relevant statistics. Regarding the molecular descriptors (MD), we used the DRAGON 7.0® software to calculate the variables that have been used for the development of the models. This software suite is able to calculate up to 7500 different descriptors, belonging to very different classes, such as topological, two-dimensional (2D), three-dimensional (3D), connectivity, and so on [54]. In order to select the best subset of MD, we have performed a feature selection process using a forward stepwise methodology [35] for both models. At the end of this procedure, we were able to develop the pristine and the carboxylate model using respectively two and three MD belonging to the topological class. The two MD used in the SWCNT-pristine model are the Narumi geometric topological index (GNAR) and the electro-topological positive variation (MAXDP). The Narumi index of a graph *G* is defined as the product of the degrees of all its vertices:

$$\mathrm{NK}(G) = \prod_{i=1}^{n} d_G(v_i) \qquad (6)$$

The MAXDP is calculated as follows:

$$S_i = I_i + \quad \Delta I_{i=} \; I_i + \quad \sum_{J0i}^{nSK} \frac{I_i - I_j}{\left(d_{ij} + 1\right)^k} \qquad (7)$$

which is calculated as the maximum positive value of ΔIi.

Regarding the SWCNT-COOH model, the continuous model was developed using three MD, one is the same GNAR used for the pristine model. The other two are defined as follows: The first one is the path/walk Randic shape indices that are calculated by summing, over the non-H atoms, the ratios of the atomic path count over the atomic walk count of the same order k and then, dividing by the total number of non-H atoms (nSK). Since path/walk count ratio is independent of molecular size, these descriptors can be considered as measures of molecular shape. Dragon calculates path/walk shape indices from order 2 up to 5, and the index of first order is not provided as the counts of the paths

and walks of length one are equal and, therefore, the corresponding molecular index equals one for all molecules. The formula in this case is not reported in the Dragon manual.

Finally, the last molecular descriptor used is the so-called lopping centric index (LOC), which is calculated as the mean information content derived from the pruning partition of a graph:

$$LOC = \sum k \quad \frac{n_k}{nSK} * \quad log_2 \frac{n_k}{nSK} \qquad (8)$$

where nk is the number of terminal vertices removed at the *k*th step and nSK is the number of non-H atoms. All the information regarding the descriptors employed in the nano-QSTR models can be retrieved from the Dragon webpage (https://chm.kode-solutions.net/products_dragon_descriptors.php).

8.3 Results and Discussion

8.3.1. CNT Effects on Submitochondrial Particles (SMP)

Herein, we present the in vitro assay on the inhibitory effect of the SWCNT ligands (SWCNT-pristine, SWCNT-COOH) at the range of concentration of 0.5–5 µg/mL over F0-ATPase using isolated rat liver submitochondrial particles (isolated F0F1-ATPase) from mitochondrial inner membrane. In general, we can see that the tested SWCNT exhibit high ability to act as F0-ATPase inhibitors (ATP-hydrolysis) at a range of concentration of 3–5 µg/mL. Besides, a concentration dependence with significant statistical difference ($p < 0.05$) when compared with SMP (untreated SMP group) and the DMSO-treated SMP was observed. We note an oligomycin A-like pattern (positive control group used) for both SWCNT ligands in a range of concentration of 3–5 µg/mL without significant statistical difference ($p > 0.05$) when compared with oligomycin A (**Figure 1**). According to this, the treated SMP from mixed CNT ligand (5 µg/mL) plus oligomycin A (1 µM) showed the strongest F0-ATPase inhibition ($p < 0.05$) when compared with untreated SMP and the DMSO-treated SMP, and the remaining CNT-treated SMP (3–5 µg/mL). This may suggest a strong synergistic effect on F0-ATPase inhibition (mitochondrial nanotoxicity). Details of these experimental results can be seen in **Figure 1**.

8.3.2. Modeling F0ATPase Inhibition Induced by SWCNTs

Herein, molecular docking was carried out in order to evaluate the influence of the carbon nanotubes (SWCNT-pristine and SWCNT-COOH) on the F0-ATPase inhibition response. The best docking binding pose from each modeled CNT (SWCNT-pristine, SWCNT-COOH) theoretically suggests that these CNT could act in the same biophysical environment as the oligomycin A based on hydrophobic non-covalent interaction (π-π interactions) involving phenylalanine hydrophobic residues (Phe 55 and Phe 64 of the chains C, D, and M), which are critically involved in the F0-ATPase inhibition (ATP-hydrolysis) in the F0-ATPase subunit active binding site, see **Figure 2**.

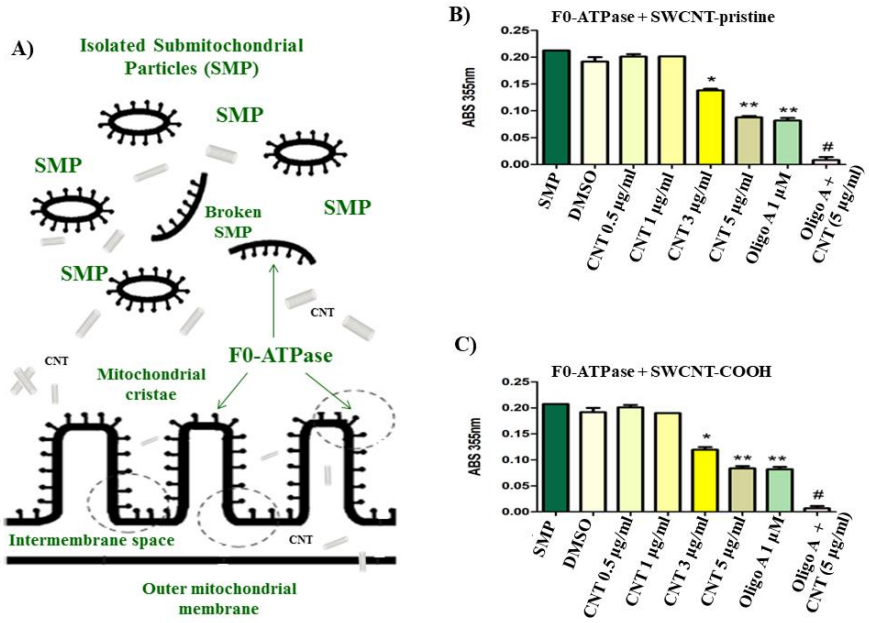


Figure 2. (**A**) Schematic representation of carbon nanotubes (CNT) interacting with isolated rat liver submitochondrial particles (SMP as F0-ATPase). (**B**, **C**) Results of experimental in vitro evaluation of the F0-ATPase inhibition induced by CNT (i.e., CNT as SWCNT-pristine or SWCNT-COOH) using F0-ATPase under the different conditions described in the Material and Methods Section as treatments: (i) untreated submitochondrial particles control (SMP as F0-ATPase), (ii) DMSO-treated SMP, (iii) CNT-treated SMP (1–5 µg/mL), (iv) oligomycin A-treated SMP (oligomycin A is a specific F0F1-ATPase inhibitor used as a positive control), and (v) treated SMP mixed with SWCNT or SWCNT-COOH at concentration of 5 µg/mL + oligomycin A (1 µM) to mimick synergistic effects on F0-ATPase inhibition, which was performed as an additional control group. Results are representative of three experiments (n = 3). Symbols (*, **, #) were used to denote statistical differences ($p < 0.05$) between the evaluated experimental

groups used in the in vitro assay containing the SMP. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

The free energy of binding (FEB) values of the formed docking complexes follow the order: FEB (oligomycin A/F0-ATPase complex) = −9.8 kcal/mol > FEB (SWCNT-COOH/F0-ATPase complex) = −6.8 kcal/mol ~ FEB (SWCNT-pristine complex) = −5.9 kcal/mol, with interatomic distance of interaction lower than 5 Å, in all the cases. Besides, we note the presence of π-π interactions, like Y-shaped and pseudo parallel-displaced motif-orientation preferences, for both single-walled carbon nanotubes. Besides, more electrostatically favored interactions in the CNT-sidewall than the CNT-tips were observed in both simulations (SWCNT-pristine and SWCNT-COOH). This was probably due to better orientation and stability between the planar-benzene-quadrupoles formed between van der Waals surface from the modeled SWCNT and the phenylalanine hydrophobic residues (Phe 55 and Phe 64) of the F0-ATPase binding site and interacting in the same biophysical environment as the F0-ATPase-specific inhibitor (oligomycin A) [17]. Please, see **Figure 3.**

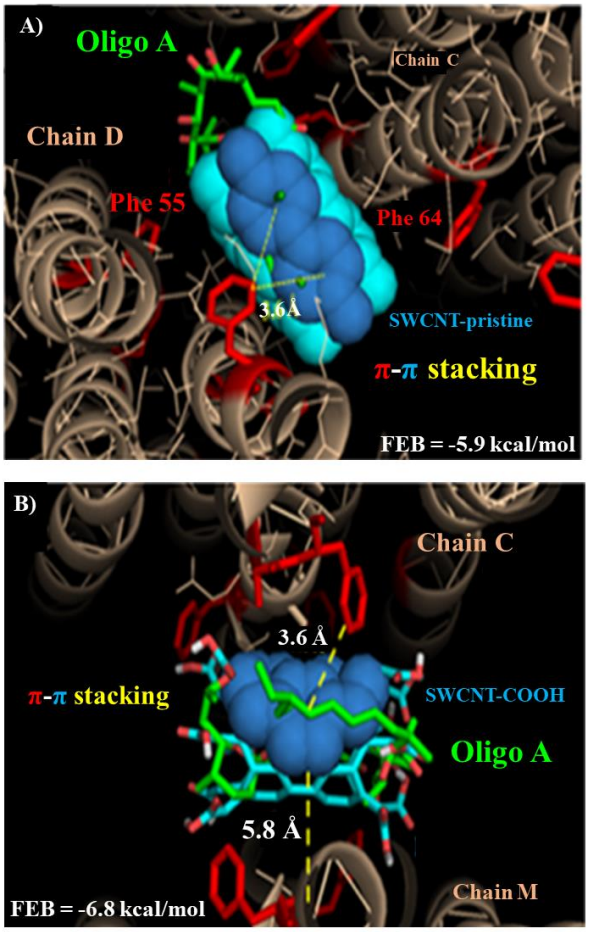


Figure 3. Snapshots selection from molecular docking interactions obtained from the best binding poses of the ligands as (**A**) superimposed representation of oligomycin A and SWCNT-pristine, and (**B**) superimposed representation of oligomycin and SWCNT-pristine and SWCNT-COOH interacting with critical phenylalanine hydrophobic residues (Phe 55 and Phe 64: labelled red) which belong to the target chains C, D, and M in the F0-ATPase subunit receptor. Please note that oligomycin A (labelled green) corresponds to the control simulation experiment used here as a reference due to this ligand being the specific inhibitor of the F0-ATPase in all cases. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

Next, we carried out the theoretical modeling based on the local perturbation response scanning maps (LPRS maps). The LPRS maps are based on elastic network models (ENM models) and have been widely recognized to study relevant conformational changes promoted from distance-based fluctuations in the alpha carbons (C(α)) of a given target protein (as F0-ATPase under unbound and bound states) at the atomistic and molecular level [53]. It is well-known that the ENM models could explain a large number of the conformational differences based on the perturbation patterns of the network formed by the target residues evaluated (Phe 55 and Phe 64). In this instance, LPRS maps generate comprehensive visualizations of the F0-ATPase inhibition response, which allows to evaluate allosteric signal propagations in response to external perturbations under the presence of a given ligand (i.e., the oligomycin A as a F0-ATPase-specific inhibitor, SWCNT-pristine, and SWCNT-COOH). The results can be seen in **Figure 4**.

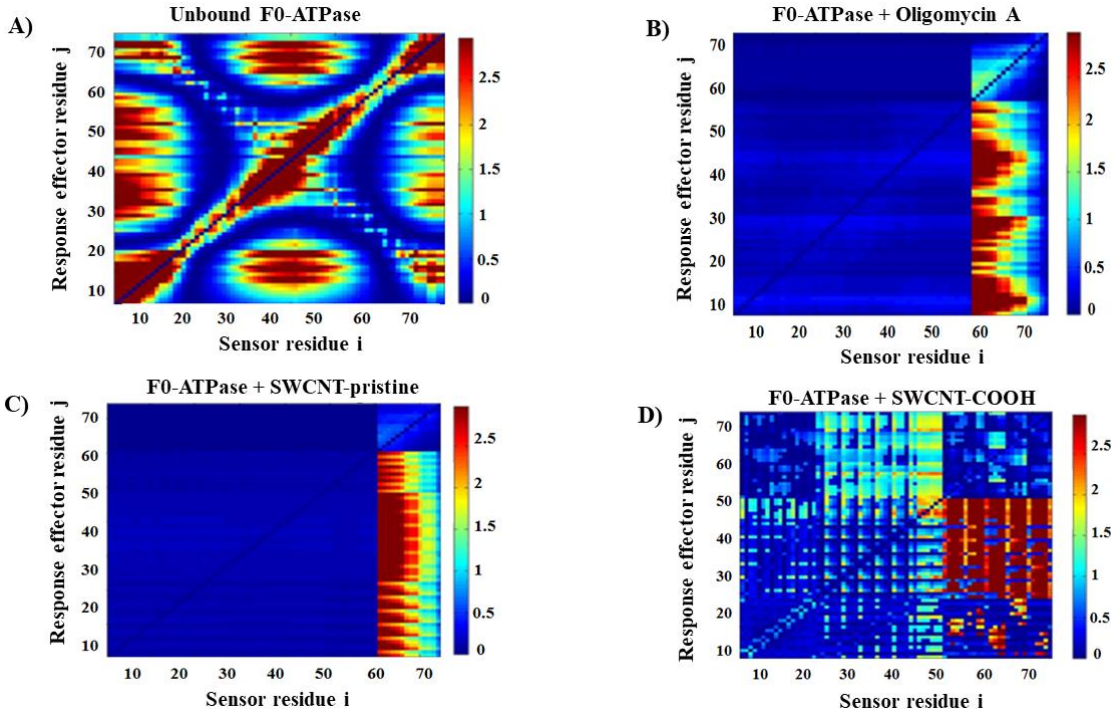


Figure 4. Perturbation response analysis for the F0-ATPase inhibition response. (**A**) LPRS map obtained for the unbound F0-ATPase as the control simulation experiment. Individual LPRS maps obtained from the best docking complexes (in the bound state for all the ligands tested) with intensity bar color representing the i,j residue perturbations (on the right) for: (**B**) oligomycin A/F0-ATPase complex, (**C**) SWCNT-pristine/F0-ATPase complex, and (**D**) SWCNT-COOH/F0-ATPase complex. All the LPRS maps were established in range of the low-frequency normal modes in order to capture relevant fluctuations associated with F0-ATPase catalytic function across the different conditions simulated. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

The results on LPRS maps show that both single-walled carbon nanotubes (SWCNT-pristine and SWCNT-COOH) promote a significant change in the perturbation patterns of the network of target residues compared with the physiological condition represented by the unbound state of F0-ATPase. In this regard, we note abrupt perturbations in several blocks of residues more pronounced for the SWCNT-pristine (strong F0-ATPase inhibition) than the SWCNT-COOH (moderate F0-ATPase inhibition) during the interaction with the F0-ATPase. Interestingly, the LPRS map of the SWCNT-pristine/F0-ATPase complex mimicked the toxicodynamic behavior of the oligomycin A/F0-ATPase complex, inducing strong F0-ATPase inhibition (see Figure 4B, C), suggesting a similar pattern of allosteric network perturbation. However, the LPRS map obtained from the

SWCNT-COOH/F0-ATPase complex (Figure 4D) exhibits a pattern of perturbation less affected when compared with the physiological condition depicted for the F0F1ATPase unbound state (Figure 4A), maintaining a certain structural and functional coupling between the residues composing the F0-ATPase network, suggesting the presence of a moderate nanotoxicity-based F0-ATPase inhibition. The relevance of these results is that strong local perturbations similar to those observed in Figure 4A, B are able to induce strong allosteric perturbations in the j-effector residues from the F0-ATPase receptor, affecting its mitochondrial catalytic function (ATP-hydrolysis) involving the signal transduction of the perturbations from the block of i-sensor residues which trigger abnormal signals' propagation across the inter-residue network for j-effector F0-ATPase residues. We could suggest that considering the SWCNT docking position, both ligands (SWCNT-pristine >> SWCNT-COOH) can theoretically disrupt the H^+-proton flux dynamic in the mitochondrial H^+-F0-ATPase subunit, compromising the coupling between oxidative phosphorylation and electron transport in the respiratory chain, inducing potential bioenergetic dysfunction and the mitochondria nanotoxicity [9].

In order to quantify potential fractal geometrical perturbations, a fractal surface analysis was carried out to model changes-based perturbations in the geometric surface of the binding effector residues of the F0-ATPase under unbound and bound states (i.e., under SWCNT-pristine and SWCNT-COOH interactions) [9]. Several fractal dimensions (FDs, namely: D_{BW}, D_{B+BW}, and D_{W+BW}) were calculated using the box-counting methods from the LPRS maps previously obtained [55]. The Fractal Theory allows the mathematical modeling of the geometric complexity (across multiple scales) and self-similarity (scale-invariant structure) from non-Euclidean real or virtual objects (such as the tested SWCNT). One of the most important properties in the fractal modeling is the degree of self-similarity. Then, a topological fractal dimension near to 2 is categorized-like, high complexity (i.e., high variety of geometrical information after the docking interaction) and low self-similarity; in contrast, a topological fractal dimension closer to 1 informs about little complexity and high self-similarity after the docking interaction. Herein, the non-Euclidean geometrical patterns were included according to the fractal dimension, like FD_{BW}, that describes the surface geometric perturbations in the border of the LPRS map fractal pattern [55]. The FD_{B+BW} characterizes the surface geometric perturbations on the white background, and the FD_{W+BW} characterizes the fractal perturbations pattern on the black background from the LPRS images calculated for each simulation condition, see **Figure 5**.

Herein, the obtained FDs are related to the F0-ATPase surface and backbone non-Euclidean geometry [9,55]. FDs inform about how the F0-ATPase folding, packing density, solvent accessibility, and binding interaction properties could be perturbed under the presence of different ligands forming docking complexes (oligomycin A/F0-ATPase complex, SWCNT-pristine/F0-ATPase complex, and SWCNT-COOH/F0-ATPase complex). In this context, we suggest that, in the bound state (i.e., during the docking interaction), the SWCNT-pristine led to higher F0F1-ATPase nanotoxicity-based allosteric perturbations than its carboxylate analogous (SWCNT-COOH) based on their obtained values for the fractal dimensions (FD_{BW}), such as SWCNT-pristine/F0-ATPase complex (FD_{BW} = 1.29) < SWCNT-COOH/F0-ATPase complex (FD_{BW} = 1.45), which quantitatively exhibits very similar features-based fractal dimension (D_{BW}, D_{BBW}, and D_{WBW}) compared to physiological condition (unbound F0-ATPase (FD_{BW} = 1.45)) used as a control for comparison purposes. It is well-known that slight variations in the fractal dimension as observed in the bound state for the docking complexes SWCNT-pristine/F0-ATPase and SWCNT-COOH/F0-ATPase (Figure 5C, D, respectively) are sufficient to induce changes in the geometry and roughness of the active site of F0-subunit of the F0F1-ATPase.

These results fit well with the previous LPRS maps, strongly suggesting that the SWCNT-pristine/F0-ATPase complex (FD_{BW} = 1.29) mimicked the nanotoxicological behavior of the specific F0F1-ATPase inhibitor (oligomycin A) with very close calculated fractal dimension for oligomycin A/F0-ATPase complex (FD_{BW} = 1.32), both lower than the physiological condition of unbound F0-ATPase cited above (Figure 5A). As previously cited, a FD ≈ 2 reveals a high variety of geometrical information and low self-similarity, while FD ≈ 1 represents little complexity and high self-similarity. On the other hand, the FD values obtained for FD_{B+BW} and FD_{W+BW} remain as unperturbed around 1.85 in all the cases, revealing high complexity of geometrical information [9,55].

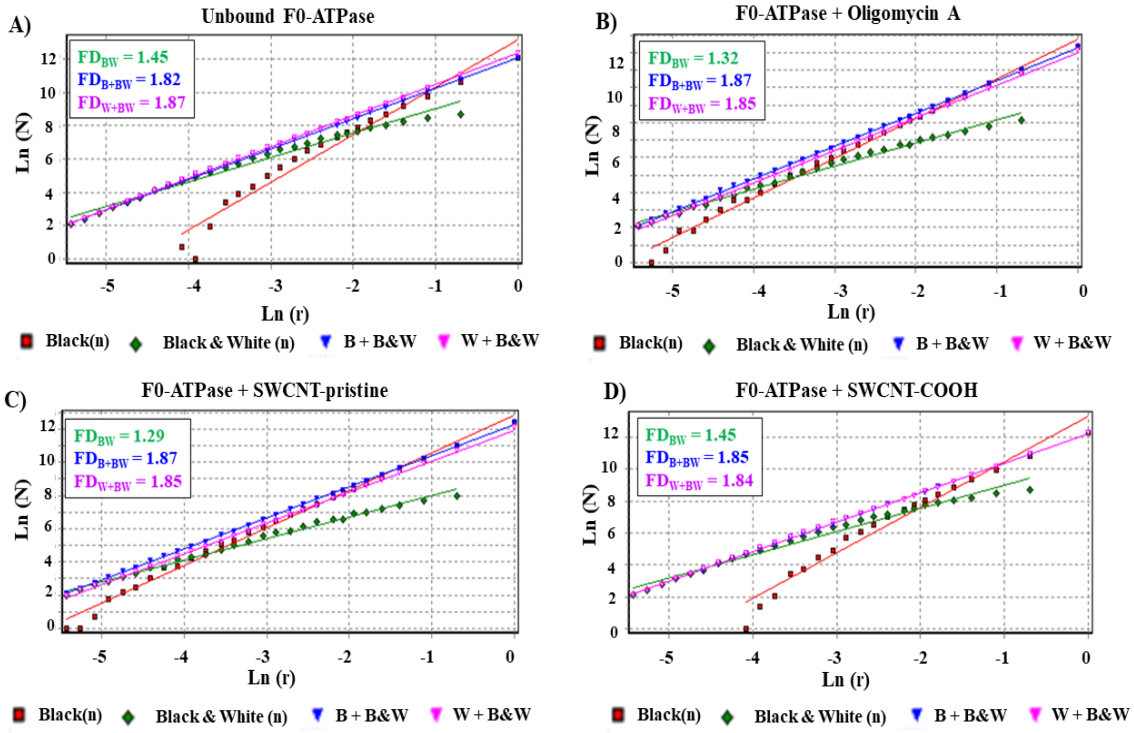


Figure 5. Fractal spectrum based on the box-counting method performed to obtain the slopes of the linear regression yields from binary black/white LPRS maps image-processing. These slopes represent the fractal dimensions (FD: D_{BW}, D_{B+BW}, and D_{W+BW}) for the best docking complexes, namely: (**A**) unbound F0-ATPase, (**B**) oligomycin A/F0-ATPase complex, (**C**) SWCNT-pristine/F0-ATPase complex, and (**D**) SWCNT-COOH/F0-ATPase complex. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

The results of fractal surface perturbation suggest that the SWCNT-pristine can induce significant changes in the geometrical selectivity of the F0-ATPase, like oligomycin A. It is well-known that perturbation (global and local perturbations) in the three-dimensional spatial arrangement of atoms composing effector residues (*j*-effector allosteric residues) of proteins can be studied using their FDs. Fractal surface perturbations could negatively impact on catalytic function of F0-ATPase, irreversibly affecting the structural properties of the binding cavities, which are of paramount importance in the complementary processes like substrate recognition and ligand geometrical specificity. Probably, topologically perturbed van der Waals fractal surface of F0-ATPase after the docking interaction with SWCNT-COOH could theoretically explain the moderate mitochondrial nanotoxicity observed from the SWCNT-COOH/F0-ATPase docking complex (refer to **Figure**

4A, **D**). Lastly, we carried out a nano-quantitative structure–toxicity relationship approach (Nano-QSRT models) in order to evaluate the influence of additional geometric properties of the ligands SWCNT-pristine and SWCNT-COOH based on the well-known relationship between the topology geometry based on the n, m Hamada index with their nanotoxicological properties (i.e., SWCNT-mitotoxicity).

8.3.3. Performed Nano-QSTR Models

As reported in the Material and Methods Section, the Nano-QSTR model for SWCNT-pristine was developed using only two variables belonging to the topological index category. The observed versus predicted values and the other relevant statistics are reported in the Tables 1 and 2, and **Figure 6**, respectively. In addition, we have also reported the AD in **Figure 7**.

Table 1. Results of the Nano-QSTR regression model for mitochondrial F0-ATPase inhibition induced by SWCNT-pristine. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

SWCNT-Pristine (n, m)	Data Observed	Data Predicted	Data Residuals	Cross-Validation [a, b]
amchair 3.3	−20.00000	−18.93350	−1.06650	training
amchair 4.4	−19.70000	−18.83954	−0.86046	training
amchair 5.5	−18.80000	−18.77444	−0.02556	training
amchair 6.6	−18.50000	−18.72592	0.22592	validation
amchair 7.7	−18.20000	−18.68908	0.48908	training
amchair 8.8	−17.50000	−18.66083	1.16083	training
amchair 9.9	−17.20000	−18.63872	1.43872	training
chiral 3.2	−17.20000	−16.28865	−0.91135	validation
chiral 4.1	−17.20000	−15.58908	−1.61092	training
chiral 4.2	−17.00000	−15.84427	−1.15573	training
chiral 4.3	−16.30000	−15.96788	−0.33212	training
chiral 5.1	−16.20000	−15.56891	−0.63109	validation
chiral 5.2	−16.20000	−15.44925	−0.75075	training
chiral 5.3	−16.00000	−15.63349	−0.36651	training
chiral 5.4	−16.00000	−15.72809	−0.27191	training
chiral 6.1	−16.00000	−15.26864	−0.73136	validation
chiral 6.2	−15.90000	−15.19863	−0.70137	training
chiral 6.3	−15.90000	−15.37446	−0.52554	training
chiral 6.4	−15.90000	−15.47126	−0.42874	training
chiral 6.5	−15.80000	−15.39773	−0.40227	validation
chiral 7.1	−15.70000	−15.02346	−0.67654	training
chiral 7.2	−15.40000	−15.17288	−0.22712	training
chiral 7.3	−15.40000	−15.10112	−0.29888	training
chiral 7.4	−15.20000	−15.24345	0.04345	validation
chiral 7.5	−15.20000	−15.34078	0.14078	training
chiral 7.6	−15.20000	−15.28875	0.08875	training
chiral 8.1	−15.00000	−15.03210	0.03210	training
chiral 8.2	−15.00000	−14.77537	−0.22463	validation
chiral 8.3	−14.90000	−15.23422	0.33422	training
chiral 8.4	−14.80000	−15.08583	0.28583	training
chiral 8.5	−14.70000	−15.15833	0.45833	training
chiral 8.6	−14.70000	−15.24567	0.54567	validation

chiral 8.7	−14.70000	−15.20136	0.50136	training
chiral 9.3	−14.60000	−14.79209	0.19209	training
chiral 9.4	−14.50000	−15.12489	0.62489	training
chiral 9.5	−14.50000	−15.11612	0.61612	validation
chiral 9.6	−14.50000	−15.09848	0.59848	training
chiral 9.7	−14.30000	−15.29203	0.99203	training
zigzag 3.0	−14.30000	−15.37860	1.07860	training
zigzag 4.0	−14.10000	−14.00654	−0.09346	validation
zigzag 5.0	−13.70000	−14.00654	0.30654	training
zigzag 6.0	−13.70000	−13.79221	0.09221	training
zigzag 7.0	−13.60000	−13.62087	0.02087	training
zigzag 8.0	−13.30000	−13.48024	0.18024	validation
zigzag 9.0	−12.90000	−13.36294	0.46294	training

Datasets: training [a] and validation sets [b].

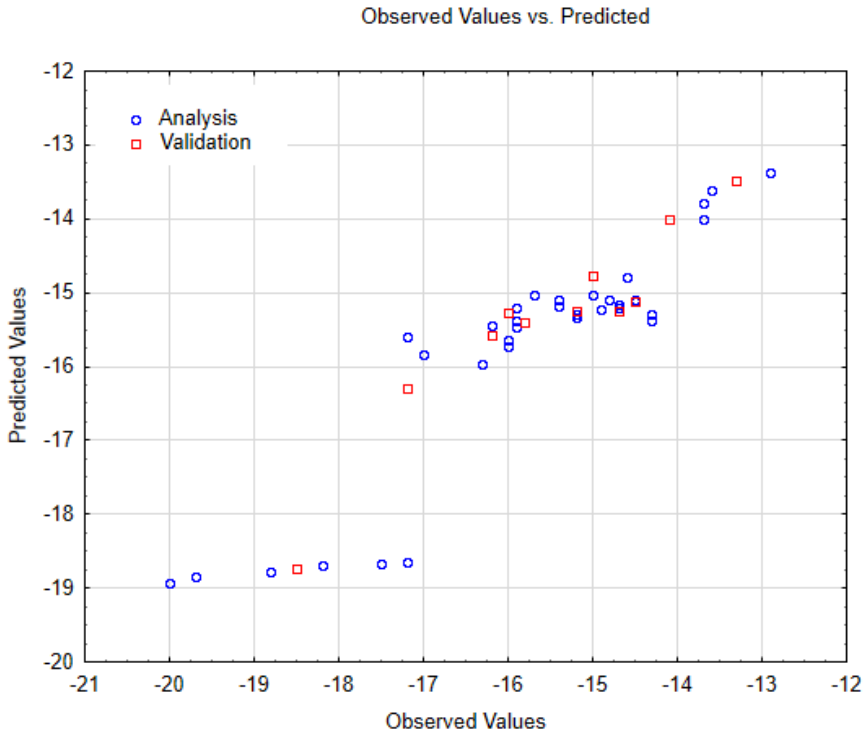


Figure 6. Results of observed versus predicted values obtained for the Nano-QSTR regression model performed for the SWCNT-pristine data. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

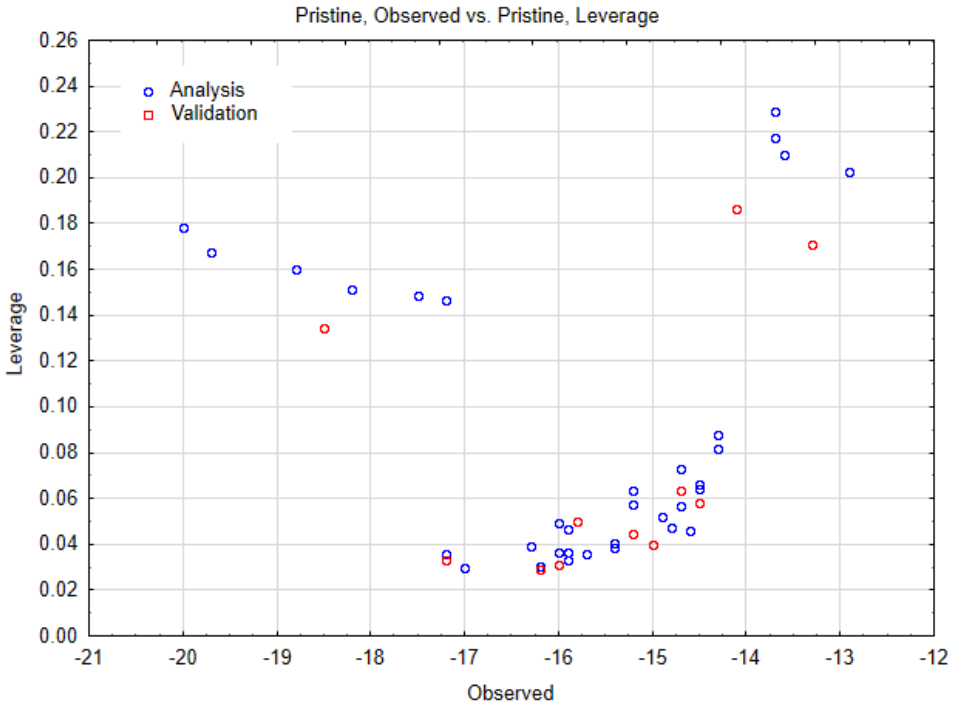


Figure 7. Applicability domain for SWCNT-pristine data. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

Table 2. Results of the relevant statistical parameters obtained from the Nano-QSTR regression model for SWCNT-pristine. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

Statistical Parameters	Value
Multiple R	0.911445
Multiple R^2	0.830731
Adjusted R^2	0.819811
Sum of squares Model	77.70196
Degrees of freedom Model	2
Mean squared errors Model	38.85098
Sum of squares Residual	15.83245
Degrees of freedom Residual	31
Mean squared errors Residual	0.510724
F	76.07035
P	0

As can be seen in the Tables 1 and 2, the Nano-QSTR model shows an overall high accuracy and goodness of fit, thus indicating that this model can be used for a continuous prediction of the likelihood of induced mitochondria nanotoxicity inhibition on F0F1-ATPase by interaction with SWCNT-pristine (f(FEB_1)). In this regard, the best Nano-QSTR regression model is based on the linear **Equation (9)** as:

$$f(FEB\_1) = -8.24425(GNar) + 0.614121(MAXDP) - 2.87142 \quad (9)$$

Afterward, we performed a Nano-QSTR model for SWCNT-COOH. For this instance, was carried out a QSTR regression model by using three variables, as in the case of the previous model (i.e., using SWCNT-pristine). Herein, the results obtained on observed versus predicted values, and the other relevant statistical parameters, are summarized in the **Tables 3** and **4**, and **Figure 8**, respectively. In addition, we have also reported the AD in **Figure 9.**

Table 3. Results of the Nano-QSTR regression model for mitochondrial F0-ATPase inhibition induced by SWCNT-COOH. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

SWCNT-COOH (n, m)	Data Observed	Data Predicted	Data Residuals	Cross-Validation [a, b]
amchair 3.3	−34.80000	−33.04305	−1.75695	training
amchair 4.4	−33.10000	−31.92664	−1.17336	training
amchair 5.5	−32.30000	−31.36844	−0.93156	training
amchair 6.6	−32.30000	−30.81023	−1.48977	validation
amchair 7.7	−29.80000	−30.53113	0.73113	training
amchair 8.8	−29.10000	−30.25203	1.15203	training
amchair 9.9	−29.00000	−29.97293	0.97293	training
chiral 3.2	−28.50000	−26.92826	−1.57174	validation
chiral 4.1	−27.90000	−26.16358	−1.73642	training
chiral 4.2	−27.60000	−27.01059	−0.58941	training
chiral 4.3	−26.90000	−25.00316	−1.89684	training
chiral 5.1	−26.70000	−28.51704	1.81704	validation
chiral 5.2	−26.60000	−24.71436	−1.88564	training
chiral 5.3	−26.40000	−25.08549	−1.31451	training
chiral 5.4	−26.40000	−24.19447	−2.20553	training
chiral 6.1	−26.30000	−27.20385	0.90385	validation
chiral 6.2	−25.90000	−24.54620	−1.35380	training
chiral 6.3	−25.40000	−25.47553	0.07553	training
chiral 6.4	−25.00000	−24.13724	−0.86276	training
chiral 6.5	−24.80000	−24.44496	−0.35504	validation
chiral 7.1	−24.70000	−25.21534	0.51534	training
chiral 7.2	−24.60000	−26.28422	1.68422	training
chiral 7.3	−24.50000	−24.04522	−0.45478	training
chiral 7.4	−24.30000	−24.52728	0.22728	validation
chiral 7.5	−24.30000	−23.88676	−0.41324	training
chiral 7.6	−24.30000	−23.32855	−0.97145	training
chiral 8.1	−24.10000	−26.95336	2.85336	training
chiral 8.2	−24.10000	−23.51211	−0.58789	validation
chiral 8.3	−24.10000	−25.39320	1.29320	training
chiral 8.4	−24.00000	−24.63822	0.63822	training
chiral 8.5	−23.70000	−24.27680	0.57680	training
chiral 8.6	−23.50000	−22.90990	−0.59010	validation
chiral 8.7	−23.50000	−24.89841	1.39841	training
chiral 9.3	−23.00000	−23.42979	0.42979	training
chiral 9.4	−22.60000	−24.13724	1.53724	training
chiral 9.5	−22.50000	−23.71859	1.21859	validation
chiral 9.6	−22.40000	−23.16039	0.76039	training
chiral 9.7	−22.20000	−23.53151	1.33151	training
zigzag 3.0	−22.10000	−25.42752	3.32752	training

zigzag 4.0	−21.70000	−20.86901	−0.83099	validation
zigzag 5.0	−21.50000	−20.86901	−0.63099	training
zigzag 6.0	−21.40000	−20.17126	−1.22874	training
zigzag 7.0	−21.10000	−19.61305	−1.48695	training
zigzag 8.0	−20.90000	−19.19440	−1.70560	validation
zigzag 9.0	−17.30000	−18.91530	1.61530	training

Datasets: training [a] and validation sets [b].

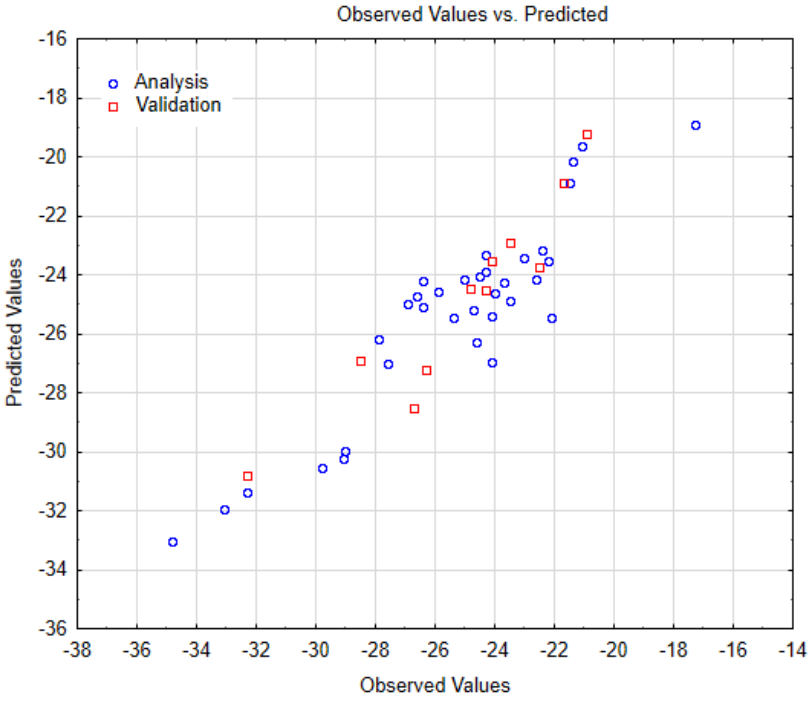


Figure 8. Results of observed versus predicted values obtained for the Nano-QSTR regression model performed for the SWCNT-COOH data. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

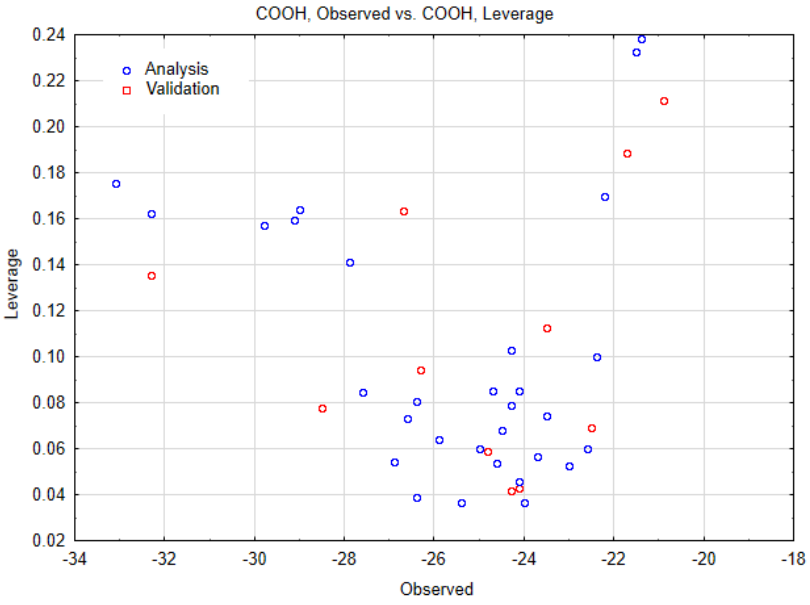


Figure 9. Applicability domain for SWCNT-COOH data. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

Table 4. Results of the relevant statistical parameters obtained from the Nano-QSTR regression model for SWCNT-COOH. Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

Statistical Parameters	Value
Multiple R	0.918915
Multiple R^2	0.844404
Adjusted R^2	0.828845
Sum of squares Model	366.1187
Degrees of freedom Model	3
Mean squared errors Model	122.0396
Sum of squares Residual	67.46364
Degrees of freedom Residual	30
Mean squared errors Residual	2.248788
F	54.26905
p	0.000000

For the case of the SWCNT-COOH dataset, the final Nano-QSTR regression model to predict the mitochondrial F0-ATPase inhibition *(f(FEB_2))* is represented by the linear **Equation (10)** as:

$$\mathrm{f(FEB\_2)} = -1005.47(GNar) - 1401.69(PW5) - 139.55(LOC) - 2326.4 \quad (10)$$

Overall, the proposed methodologies rigorously obey the Organization for Economic Co-operation and Development (OECD) and the International Organization for Standardization guidelines for development of alternative methods for Computational Nanotoxicology [56].

References

1. Fillingame, R.H.; Jiang, W.; Dmitriev, O.Y. Coupling H(+) transport to rotary catalysis in F-type ATP synthases: Structure and organization of the transmembrane rotary motor. *J. Exp. Biol.* **2000**, *203*, 9–17.
2. Mitchell, P. Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic type of Mechanism. *Nature* **1961**, *191*, 144–148.
3. Bulygin, V.V.; Duncan, T.M.; Cross, R.L. Rotor/Stator interactions of the epsilon subunit in Escherichia coli ATP synthase and implications for enzyme regulation. *J. Biol. Chem.* **2004**, *279*, 35616–35621.

4. Martínez-Cano, E.; Ortiz-Genaro, G.; Pacheco-Moisés, F.; Macías-Islas, M.A.; Sánchez-Nieto, S.; Rosales-Corral, S.A. Functional disorders of FOF1-ATPase in submitochondrial particles obtained from platelets of patients with a diagnosis of probable Alzheimer's disease. *Rev. Neurol.* **2005**, *40*, 81–85.
5. Senior, A.E. The structure of mitochondrial ATPase. *Biochim. Biophys. Acta (BBA) Rev. Bioenerg.* **1973**, *301*, 249–277.
6. Seyfried, T.N.; Shelton, L.M., Cancer as a metabolic disease. *Nutr. Metab.* **2010**, *7*, 7.
7. Yang, Z.; Zhang, Y.; Yang, Y.; Sun, L.; Han, D.; Li, H.; Wang, C. Pharmacological and toxicological target organelles and safe use of single-walled carbon nanotubes as drug carriers in treating Alzheimer disease. *Nanomed. Nanotechnol. Biol. Med.* **2010**, *6*, 427–441.
8. González-Durruthy, M.; Giri, A.K.; Moreira, I.; Concu, R.; Melo, A.; Ruso, J.M.; Cordeiro, M.N.D.S. Computational modeling on mitochondrial channel nanotoxicity. *Nano Today* **2020**, *34*, 100913.
9. González-Durruthy, M.; Manske Nunes, S.; Ventura-Lima, J.; Gelesky, M.A.; González-Díaz, H.; Monserrat, J.M.; Concu, R.; Cordeiro, M.N.D.S. MitoTarget Modeling Using ANN-Classification Models Based on Fractal SEM Nano-Descriptors: Carbon Nanotubes as Mitochondrial F0F1-ATPase Inhibitors. *J. Chem. Inf. Modeling* **2019**, *59*, 86–97.
10. González-Durruthy, M.; Castro, M.; Nunes, S.M.; Ventura-Lima, J.; Alberici, L.C.; Naal, Z.; Atique-Sawazaki, D.T.; Curti, C.; Ruas, C.P.; Gelesky, M.A.; et al. QSPR/QSAR-based Perturbation Theory approach and mechanistic electrochemical assays on carbon nanotubes with optimal properties against mitochondrial Fenton reaction experimentally induced by Fe2+-overload. *Carbon* **2017**, *115*, 312–330.
11. Atwal, K.S.; Wang, P.; Rogers, W.L.; Sleph, P.; Monshizadegan, H.; Ferrara, F.N.; Traeger, S.; Green, D.W.; Grover, G.J. Small molecule mitochondrial F1F0 ATPase hydrolase inhibitors as cardioprotective agents. Identification of 4-(N-arylimidazole)-substituted benzopyran derivatives as selective hydrolase inhibitors. *J. Med. Chem.* **2004**, *47*, 1081–1084.
12. Liu, J.; Yuan, B.; Wu, X.; Li, J.; Han, F.; Dou, Y.; Chen, M.; Yang, Z.; Yang, K.; Ma, Y. Modulated enhancement in ion transport through carbon nanotubes by lipid decoration. *Carbon* **2017**, *111*, 459–466.
13. González-Durruthy, M.; Werhli, A.V.; Seus, V.; Machado, K.S.; Pazos, A.; Munteanu, C.R.; González-Díaz, H.; Monserrat, J.M. Decrypting Strong and Weak Single-Walled Carbon Nanotubes Interactions with Mitochondrial Voltage-Dependent Anion Channels Using Molecular Docking and Perturbation Theory. *Sci. Rep.* **2017**, *7*, 13271.

14. González-Durruthy, M.; Werhli, A.V.; Cornetet, L.; Machado, K.S.; González-Díaz, H.; Wasiliesky, W.; Ruas, C.P.; Gelesky, M.A.; Monserrat, J.M. Predicting the binding properties of single walled carbon nanotubes (SWCNT) with an ADP/ATP mitochondrial carrier using molecular docking, chemoinformatics, and nano-QSBR perturbation theory. *RSC Adv.* **2016**, *6*, 58680–58693.
15. Halestrap, A.P. The C Ring of the F1Fo ATP Synthase Forms the Mitochondrial Permeability Transition Pore: A Critical Appraisal. *Front. Oncol.* **2014**, *4*, 234.
16. Alavian, K.N.; Beutner, G.; Lazrove, E.; Sacchetti, S.; Park, H.-A.; Licznerski, P.; Li, H.; Nabili, P.; Hockensmith, K.; Graham, M.; et al. An uncoupling channel within the c-subunit ring of the F_1F_OATP synthase is the mitochondrial permeability transition pore. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10580–10585.
17. Shchepina, L.A.; Pletjushkina, O.Y.; Avetisyan, A.V.; Bakeeva, L.E.; Fetisova, E.K.; Izyumov, D.S.; Saprunova, V.B.; Vyssokikh, M.Y.; Chernyak, B.V.; Skulachev, V.P. Oligomycin, inhibitor of the F0 part of H+-ATP-synthase, suppresses the TNF-induced apoptosis. *Oncogene* **2002**, *21*, 8149–8157.
18. Ponnalagu, D.; Singh, H. Insights into the Role of Mitochondrial Ion Channels in Inflammatory Response. *Front. Physiol.* **2020**, *11*, 258.
19. Missiroli, S.; Patergnani, S.; Caroccia, N.; Pedriali, G.; Perrone, M.; Previati, M.; Wieckowski, M.R.; Giorgi, C. Mitochondria-associated membranes (MAMs) and inflammation. *Cell Death Dis.* **2018**, *9*, 329.
20. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
21. Forli, S.; Huey, R.; Pique, M.E.; Sanner, M.F.; Goodsell, D.S.; Olson, A.J. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **2016**, *11*, 905–919.
22. Lavecchia, A. Machine-learning approaches in drug discovery: Methods and applications. *Drug Discov. Today* **2015**, *20*, 318–331.
23. Chennubhotla, C.; Bahar, I. Signal Propagation in Proteins and Relation to Equilibrium Fluctuations. *PLoS Comput. Biol.* **2007**, *3*, e172.
24. Li, H.; Chen, S.; Zhao, H. Fractal structure and conformational entropy of protein chain. *Int. J. Biol. Macromol.* **1990**, *12*, 374–378.
25. Weng, G.; Bhalla, U.S.; Iyengar, R. Complexity in biological signaling systems. *Science* **1999**, *284*, 92–96.

26. Lezon, T.; Shrivastava, I.; Yang, Z.; Bahar, I. *Chapter 7 Elastic Network Models for Biomolecular Dynamics: Theory and Application to Membrane Proteins and Viruses*; World Scientific Publishing: Hanbook on Biological Networks, Singapore, Singapore, 2009.
27. Riccardi, D.; Cui, Q.; Phillips, G.N., Jr. Application of elastic network models to proteins in the crystalline state. *Biophys. J.* **2009**, *96*, 464–475.
28. Flechsig, H.; Togashi, Y. Designed Elastic Networks: Models of Complex Protein Machinery. *Int. J. Mol. Sci.* **2018**, *19*, 3152.
29. Atwal, K.S.; Ahmad, S.; Ding, C.Z.; Stein, P.D.; Lloyd, J.; Hamann, L.G.; Green, D.W.; Ferrara, F.N.; Wang, P.; Rogers, W.L.; et al. N-[1-Aryl-2-(1-imidazolo)ethyl]-guanidine derivatives as potent inhibitors of the bovine mitochondrial F1F0 ATP hydrolase. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1027–1030.
30. Zhang, S.; Golbraikh, A.; Tropsha, A. Development of Quantitative Structure−Binding Affinity Relationship Models Based on Novel Geometrical Chemical Descriptors of the Protein−Ligand Interfaces. *J. Med. Chem.* **2006**, *49*, 2713–2724.
31. Burello, E.; Worth, A.P. QSAR modeling of nanomaterials. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2011**, *3*, 298–306.
32. Fourches, D.; Pu, D.; Tassa, C.; Weissleder, R.; Shaw, S.Y.; Mumper, R.J.; Tropsha, A. Quantitative Nanostructure−Activity Relationship Modeling. *ACS Nano* **2010**, *4*, 5703–5712.
33. Buglak, A.A.; Zherdev, A.V.; Dzantiev, B.B. Nano-(Q)SAR for Cytotoxicity Prediction of Engineered Nanomaterials. *Molecules* **2019**, *24*, 4537.
34. Basant, N.; Gupta, S. Multi-target QSTR modeling for simultaneous prediction of multiple toxicity endpoints of nano-metal oxides. *Nanotoxicology* **2017**, *11*, 339–350.
35. Concu, R.; Kleandrova, V.V.; Speck-Planche, A.; Cordeiro, M. Probing the toxicity of nanoparticles: A unified in silico machine learning model based on perturbation theory. *Nanotoxicology* **2017**, *11*, 891–906.
36. Ahmadi, S. Mathematical modeling of cytotoxicity of metal oxide nanoparticles using the index of ideality correlation criteria. *Chemosphere* **2020**, *242*, 125192.
37. Mu, Y.; Wu, F.; Zhao, Q.; Ji, R.; Qie, Y.; Zhou, Y.; Hu, Y.; Pang, C.; Hristozov, D.; Giesy, J.P.; et al. Predicting toxic potencies of metal oxide nanoparticles by means of nano-QSARs. *Nanotoxicology* **2016**, *10*, 1207–1214.
38. Toropova, A.P.; Toropov, A.A.; Veselinovic, A.M.; Veselinovic, J.B.; Benfenati, E.; Leszczynska, D.; Leszczynski, J. Nano-QSAR: Model of mutagenicity of fullerene as a

mathematical function of different conditions. *Ecotoxicol. Environ. Saf.* **2016**, *124*, 32–36.

39. Toropov, A.A.; Toropova, A.P. Quasi-SMILES and nano-QFAR: United model for mutagenicity of fullerene and MWCNT under different conditions. *Chemosphere* **2015**, *139*, 18–22.
40. Toropova, A.P.; Toropov, A.A. Mutagenicity: QSAR—quasi-QSAR—nano-QSAR. *Mini Rev. Med. Chem.* **2015**, *15*, 608–621.
41. Toropova, A.P.; Toropov, A.A.; Benfenati, E.; Korenstein, R.; Leszczynska, D.; Leszczynski, J. Optimal nano-descriptors as translators of eclectic data into prediction of the cell membrane damage by means of nano metal-oxides. *Environ. Sci. Pollut. Res. Int.* **2015**, *22*, 745–757.
42. Liu, R.; Rallo, R.; Weissleder, R.; Tassa, C.; Shaw, S.; Cohen, Y. Nano-SAR development for bioactivity of nanoparticles with considerations of decision boundaries. *Small* **2013**, *9*, 1842–1852.
43. Tantra, R.; Oksel, C.; Puzyn, T.; Wang, J.; Robinson, K.N.; Wang, X.Z.; Ma, C.Y.; Wilkins, T. Nano(Q)SAR: Challenges, pitfalls and perspectives. *Nanotoxicology* **2015**, *9*, 636–642.
44. Argese, E.; Bettiol, C.; Giurin, G.; Miana, P. Quantitative structure—activity relationships for the toxicity of chlorophenols to mammalian submitochondrial particles. *Chemosphere* **1999**, *38*, 2281–2292.
45. Ferguson, S.J.; Sorgato, M.C. The phosphorylation potential generated by respiring bovine heart submitochondrial particles. *Biochem. J.* **1977**, *168*, 299–303.
46. Argese, E.; Bettiol, C.; Ghelli, A.; Todeschini, R.; Miana, P. Submitochondrial particles as toxicity biosensors of chlorophenols. *Environ. Toxicol. Chem. Int. J.* **1995**, *14*, 363–368.
47. Sorgato, M.C.; Branca, D.; Ferguson, S.J. The rate of ATP synthesis by submitochondrial particles can be independent of the magnitude of the protonmotive force. *Biochem. J.* **1980**, *188*, 945–948.
48. Bergin, S.; Sun, Z.; Streich, P.; Hamilton, J.; Coleman, J. New Solvents for Nanotubes: Approaching the Dispersibility of Surfactants. *J. Phys. Chem. C* **2010**, *114*, 231–237.
49. Pedersen, P.L.; Greenawalt, J.W.; Reynafarje, B.; Hullihen, J.; Decker, G.L.; Soper, J.W.; Bustamente, E. Preparation and characterization of mitochondria and submitochondrial particles of rat liver and liver-derived tissues. *Methods Cell Biol.* **1978**, *20*, 411–481.
50. Starkov, A.A. Measurement of mitochondrial ROS production. *Methods Mol. Biol.* **2010**, *648*, 245–55.

51. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
52. Feinstein, W.P.; Brylinski, M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. *J. Cheminform.* **2015**, *7*, 18.
53. Oliwa, T.; Shen, Y. cNMA: A framework of encounter complex-based normal mode analysis to model conformational changes in protein interactions. *Bioinformatics* **2015**, *31*, i151–i160.
54. Talete srl, K. *Dragon (Software for Molecular Descriptor Calculation)*; 7.0.10; Milan,Italy, 2017.
55. González-Durruthy, M.; Scanavachi, G.; Rial, R.; Liu, Z.; Cordeiro, M.N.D.S.; Itri, R.; Ruso, J.M. Mapping the underlying mechanisms of fibrinogen benzothiazole drug interactions using computational and experimental approaches. *Int. J. Biol. Macromol.* **2020**, *163*, 730–744.
56. OECD Principles for the Validation, for Regulatory Porposes of (Quantitative) Structure Activity Relationship Model. Available online: https://www.oecd.org/chemicalsafety/risk-assessment/37849783.pdf (accessed on 13 March 2020).

Supplementary Materials

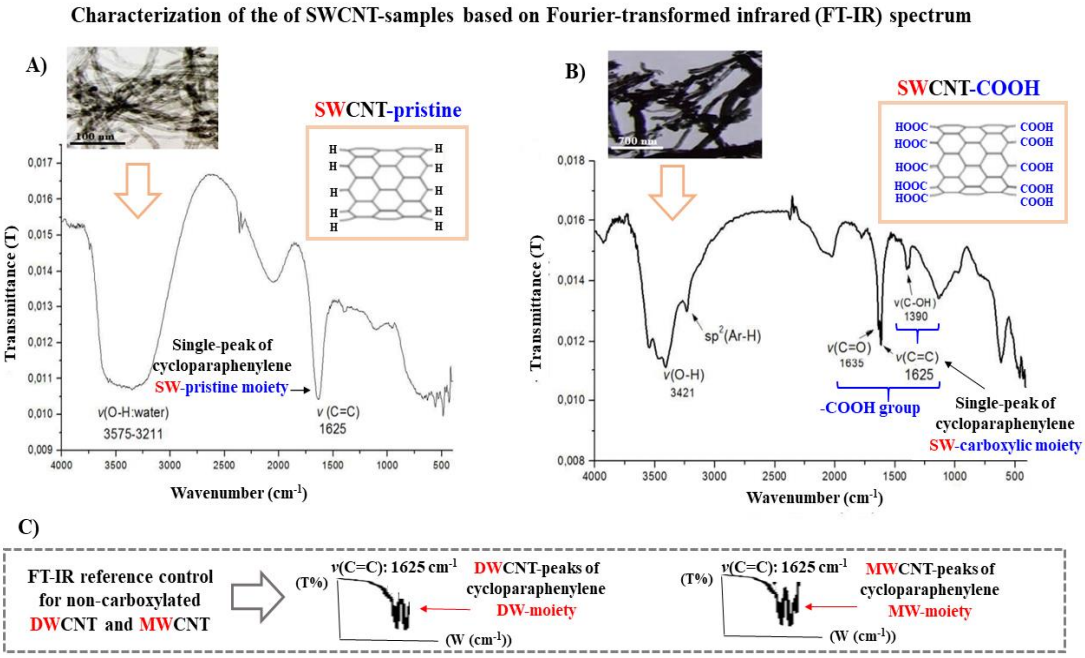


Figure S1. Characterization of the SWCNT-samples under study using Fourier-transformed infrared (FT-IR spectrum). Source MDPI article, please visit: https://doi.org/10.3390/biology10030171 (accessed on 25 February 2021)

Chapter 9.

Patricia Ferreira Schopf, Ivana Zanella da Silva, M. Natália D.S. Cordeiro, Juan M. Ruso, **Michael González-Durruthy***, Mirkos Ortiz Martins*. Nanomarker for Early Detection of Alzheimer's Disease Combining *Ab initio* DFT Simulations and Molecular Docking Approach. Biophysica 2021, 1, 76–86. https://doi.org/10.3390/biophysica1020007

Journal Impact Factor (JIF) (2021): Not available yet

CiteScore (2021): Not available yet

Author Contribution M.G.-D.: Conceptualization, methodology-based on structure-based docking, computational crystallographic validation-based Ramchandran diagrams of Alzheimer's tau protein and writing—original draft preparation. *Corresponding author

Journal authorization:

Biophysica

Biophysica is an international, peer-reviewed, open access journal that applies the approaches and methods of physics to study biological systems and is published quarterly online by MDPI.

- **Open Access** —free to download, share, and reuse content. Authors receive recognition for their contribution when the paper is reused.

Copyright and Licensing

For all articles published in MDPI journals, copyright is retained by the authors. Articles are licensed under an open access Creative Commons CC BY 4.0 license, meaning that anyone may download and read the paper for free. In addition, the article may be reused and quoted provided that the original published version is cited. These conditions allow for maximum use and exposure of the work, while ensuring that the authors receive proper credit.

In exceptional circumstances articles may be licensed differently. If you have specific condition (such as one linked to funding) that does not allow this license, please mention this to the editorial office of the journal at submission. Exceptions will be granted at the discretion of the publisher.

Chapter 9.

Nanomarker for Early Detection of Alzheimer's Disease Combining *Ab initio* DFT Simulations and Molecular Docking Approach.

Abstract

Tau protein is considered an important qualitative and quantitative biomarker for Alzheimer's disease in its asymptomatic phase. In 2011, biomarkers were suggested by the National Institute on Aging - Azheimer's Association as a new criterion for the early diagnosis of Alzheimer's disease. Thus, highlighting the non-existence of theoretical research on the subject, we investigated the binding interaction properties between phosphorylated tau protein and a theoretically modeled ligands constituted by the fullerol functionalized with radiopharmaceuticals from an *in silico* approach via molecular docking and DFT *ab initio* computational simulation. The results demonstrated that the ligand with the greatest affinity-based binding energy to the protein was fullerol+F-THK5105. However, all systems were considered promising for the development of a potential diagnostic nanomarker. These theoretical results could efficiently contribute to reduce the time and the cost for future experimental preclinical studies and open new opportunities toward molecular recognition in nanomedicine.

9.1 Introduction

Alzheimer's disease (AD) is one of the diseases that is among the trends of epidemics for the coming years, due to the increase of life expectancy of the population, since the incidence of this disease grows exponentially with age (SLOANE et al., 2002). The AD corresponds to the most common form of dementia, affecting more than 46 million people worldwide, with expectation of going up to 130 million cases by 2050 (PRINCE; BRYCE; FERRI, 2015). Despite the high worldwide incidence, only 1 in 4 people with AD are diagnosed (PRINCE et al., 2016). Statistics from 2011 revealed that approximately 36 million people had the disease in the world, but about 28 million did not receive the proper diagnostic (PRINCE; BRYCE; FERRI, 2011). These data highlight one of the main current concerns related to the disease: its differential and early diagnostic.

The symptoms of early-stage AD may be easily confused with those of other pathologies, since other diseases also have one of the main characteristics of memory lapses

(MATSUNAGA; YONEDA; 2016; GASSER; SALAMIN; ZUMBACH, 2017). Thereunto, in 2011, the National Institute on Aging – Azheimer's Association (NIA-AA) suggested the use of biomarkers as a new criterion for the identification of AD in the asymptomatic phase (JACK et al., 2011). The use of qualifying and quantifying biomarkers of biochemical changes in the brain decreases probable diagnostic errors and provides access to appropriate treatments even before clinical signs appear, minimizing the impacts caused by the disease, especially with regard to the loss of cognitive functions (BENNETT, LEIFER, 2003; SCHELTENS et al., 2016).

Tau is a protein with great preclinical relevance, because it is directly related to neuronal death, which is the main clinical sign of AD. In its normal function (i.e., physiological function), this protein is responsible for the stabilization of the microtubules, allowing the passage of current in the axons (BRUNDEN; TROJANOWSKI; LEE, 2009). The function of tau is compromised when there is a hyperphosphorylation, reducing the attachment to the microtubules and forming paired helical filament (PHF) agglomerates in the extracellular spaces, known as neurofibrillary tangles (CHIEN et al., 2014). Neurofribrillary tangles spread in the brain as AD progresses and cause synaptic loss and gradual neuronal death in neurons (MASTERS et al., 2015).

PHF-tau can be considered a promising pre-symptomatic biomarker of AD, since aggregates of this protein can be found in the brain years before the symptoms appear (CHIEN et al., 2014). Nanoscale sensors have been suggested for the detection of biological markers associated with various diseases at the cellular and molecular levels (TOUHAMI, 2015). In addition, the use of nanomaterials in the implementation of biosensors increases the speed, sensitivity and specificity of detection by the biological molecule of interest (CULLUM, VON-DINH, 2000).

Carbon-based nanomaterials are being widely used in the sensing of different diseases (AFSAHI et al., 2018; YANG et al., 2018). Particularly, the fullerol is an interesting and versatile nanomaterial derived belonging to the fullerene family. This structure can be efficiently functionalized (i.e., oxidized) with hydroxyl groups (-OH), thus making the molecule of the fullerene more hydrophilic and improving its bioavailability-based pharmacokinetics properties (BAKRY et al., 2007). Fullerene is a promising nanocarrier with respect to the passage of molecules through the blood-brain barrier (BBB), besides having excellent antioxidant and neuroprotective activity experimentally validated by in vitro experiements, improving, even, the neuronal function in AD (DUGAN et al., 1996; SANTOS et al., 2010; DA RÓZ et al., 2015). At the same time, a series of fluorescent

radiopharmaceuticals have been studied experimentally for the diagnostic of several pathophysiological processes in the brain (BENADIBA et al., 2012; VILLEMAGNE et al., 2015). These studies have demonstrated that these substances have an affinity for the PHF-tau protein and can be used as receptors of this protein in biosensors (VILLEMAGNE et al., 2015).

In this context, computational approaches could be efficiently applied for fast, cheap, and early detection of AD as an alternative strategy previous to perform experimental studies, allowing a better understanding of the molecular mechanisms of diagnostic recognition by using computational animal free-testing methodologies (ERKEKOGLU; GIRAY; BASARAN, 2011; PASSINI et al., 2017). Besides, during studies with nanostructures, the computational simulation assists in the understanding of physical and chemical phenomena occurring on the nanometric scale (VILELA NETO, PACHECO, 2012).

In this context, the main objective of the present study was to implement and obtain a prospective and computationally modeled new nanomarker formed by functionalized fullerol with nine radiopharmaceutical ligands, with potential ability on detecting early AD from the PHF-tau biomarker protein, through a combined computational approach like ab initio DFT and molecular docking simulation. To this end, the free energy of binding between the diagnostic systems (i.e., the fullerol with each radiopharmaceutical ligands) with the biomolecule (tau protein receptor) will be predicted with the corresponding molecular distances. Next, the mechanism of interaction for the best fit docking pose configuration will be studied through *ab initio* DFT simulation.

9.2 Materials and Methods

Firstly, the fullerol ($C_{60}(OH)_{24}$) was used as receptor for adding the fluorescent radiopharmaceuticals as ^{18}F-FDDNP, ^{18}F-T807, ^{18}F-T808, ^{18}F-THK523, ^{18}F-THK5105, ^{18}F-THK5117, ^{18}F-THK5351, ^{11}C-*N*-methyl Lansoprazole and ^{11}C-PBB3. The radiopharmaceutical molecules were obtained through the PubChem database (WANG et al., 2016). The immobilization of the radiopharmaceuticals to fullerol was performed through covalent bonds from the removal of water molecules from the system. Different configurations of functionalization of fullerol were analyzed with each radiopharmaceutical (nine in total). This analysis was done through *ab initio* calculations based on the Density Functional Theory (DFT) using Spanish Initiative for Electronic Simulations with Thousands of Atoms (SIESTA) computer code (SOLER et al., 2002). The exchange-correlation term in the SIESTA code was reproduced by using the Local Density Approximation (LDA)

(CEPERLEY; ALDER, 1980). The most stable fullerol-radiopharmaceutical systems were used as input ligands to address a rigid molecular docking study. Before the rigid docking execution, the potential tau protein binding-sites were predicted by using DeepSite.[28] This task was carried out here by delimiting the van der Waals surface area of the tau-protein binding residues, where the binds of the ligands (i.e., fullerol-radiopharmaceutical systems) take place. The DeepSite method (freely available at www.playmolecule.org) (JIMÉNEZ et al., 2017) is mainly based on a 3D-deep convolutional neural networks approach (DCNNs) which efficiently predict the resulting tau-protein docking box of simulation. For this instance, the Cartesian coordinates generated from the DeepSite volumetric maps of the tau-protein binding sites were used to setup a rigid-docking box simulation in the best-ranked binding site of tau protein, and considering the DeepSite maximum score and maximum volume of the selected cavity, which is well-known to have the highest druggability properties or ability to binds small ligands like fullerol-radiopharmaceutical systems. Them for the tau-protein, the details of the box-simulation are: grid box-size Cartesian coordinates of (X = 27 Å, Y = 27 Å, Z = 27 Å) with volume V= 827.02 Å^3, and grid box-center Cartesian coordinates centered at X = 192.8 Å, Y = 130 Å, Z = 142.3 Å. Afterward, all the molecular docking runs were performed using the AutoDock Vina package, with a Vina scoring function implemented by Trott and Olson. (TROTT; OLSON, 2010). Molecular docking is based on the prediction of best crystallographic binding conformation and the determination of the free energy of binding (FEB, kcal/mol) which quantify the affinity between the protein-binding complexes. To this end, the macromolecule (tau-protein), was obtained from the Protein Data Bank (PDB) website (BERMAN et al., 2000), as PDB ID: 5O3L, resolution 3,4 Å was pretreated through of the AutoDock and PyMOL softwares for the removal of crystallographic water molecules and ions.

The best crystallographic binding configuration was check for the best fitting structure (fullerol-radiopharmaceutical systems), and based on two criteria: (i) most negative affinity or FEB value from the obtained docking complexes; and (ii) the root-mean- square-deviation (R.M.S.D) for the atoms positions for the obtained docking complexes with an optimal value of R.M.S.D < 2 Å for all the simulation experiments (DURRUTHY et al., 2017). In order to provide more reliable and accurate results, the AutoDock exhaustiveness parameter equal to 50 (i.e., representing 50 fullerol-radiopharmaceutical docking-pose configurations in the modeled rigid-pocket of tau-protein) was established (TROTT; OLSON, 2010).

The specific regions of the docking interactions, as well as the type of interactions, were verified through the LigPlot program (LASKOWSKI; SWINDELLS, 2011). This program

evaluates the performance of non-covalent interactions between a protein and the ligand through a 2D diagram that shows the shape of interaction that occurred in the coupling (hydrophobic, hydrogen bonding, and electrostatics interactions) in the formed docking complexes as fullerol-radiopharmaceutical plus tau-protein. The resulting images of the atoms involved in the interactions of the radio-ligands and the atoms of the interacting amino-acid residues of tau-protein, as well as the corresponding values of the inter-atomic distances, were depicted by using PyMOL.

The interface of the post-docking system was studied through *ab initio* calculations using the SIESTA code. Then, the electronic and structural properties between the best diagnostic system composed by the fullerol-radiopharmaceutical ligand (i.e., more stable configuration obtained by DFT approach), and the most critical amino acid residue from tau protein (i.e., identified with the shortest interaction distance) were evaluated according to the binding energy of the system calculated as show the **equation 1**:

$$E_{lig} = E_{A+L} - (E_A + E_L) \quad \textbf{(1)}$$

Where E_{A+L} is the total energy (eV) of the interface of the amino acid+ligand system, E_A energy from isolated amino acid and E_L is the isolated binding energy.

Furthermore, the difference between the Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) (ΔHL) was also calculated. The isosurface used for analyzed the charges concentrations in the HOMO and LUMO was 0.001 e-/Bohr^3.

Next, the molecular structures were depicted using the VMD program (HUMPHREY; DALKE; SCHULTEN, 1996).

9.3 Results

9.3.1. Ligand modeling and preparation

The radiopharmaceutical systems (nanomarker ligands) were built by the fullerol bound to differents 9 radiopharmaceutical ligands. To this end, the initial configurations to carried out the DFT ab initio simulations were manually stablished by varying the position of the covalently attachment atoms of the radiopharmaceutical ligands to the fullerol. Then, the final interacting positions of these bonds were chosen based on distribution of charges of the systems from the isolated structures. Besides, was considered a cutoff inter-atomic distance values (dij) < 7Å as initial position of the interacting DFT configurations for the interacting atoms to build the diagnostic systems as functionalized fullerol (i-atoms) and each one of the

9 radiopharmaceutical ligands (j-atoms). This DFT configurations-based inter-atomic distance criterion (dij < 7Å) is adopted taking into account that, only atoms with $d_{ij} \leq 7$ Å are considered as relevant interacting atoms for a given binary system (or complex systems) for computational modeling purposes. Afterward, were chosen the lower energy DFT configurations in a module among those studied to built the diagnostic nanosystems formed by each one of the 9 radiopharmaceutical ligands (see **Figure 1**).

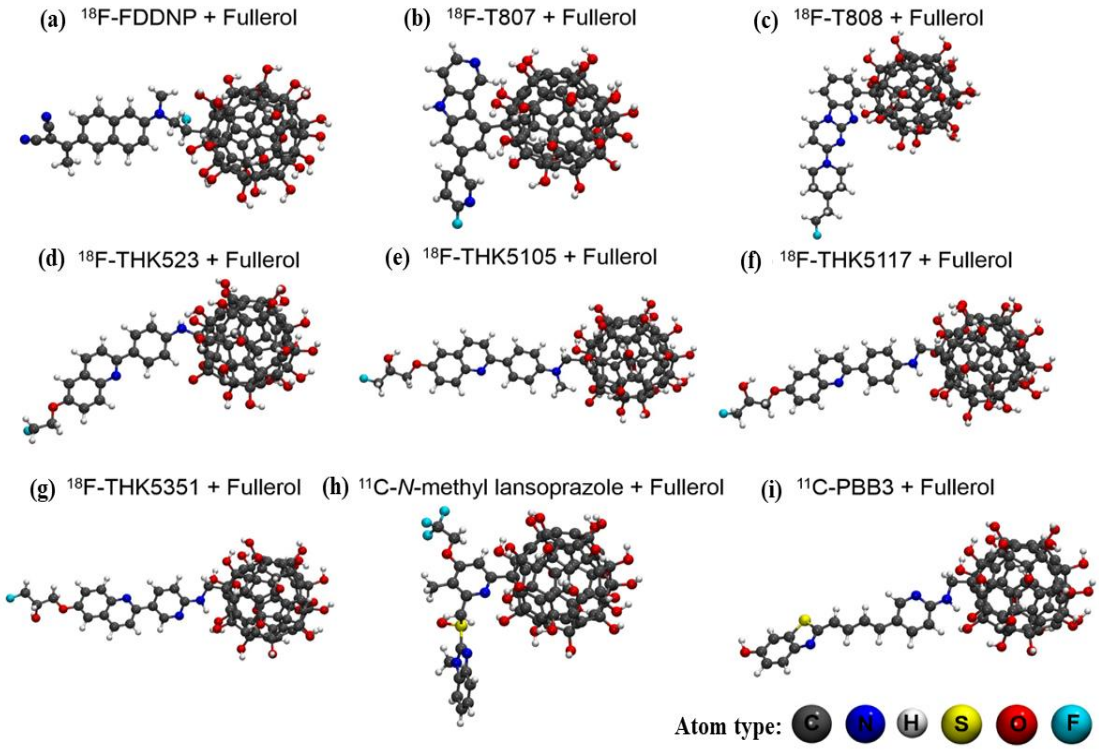


Figure 1. Optimized structures of the ligands used for the study of interaction with the PHF-tau protein. Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

9.3.2. Modeling PHF-tau protein and ligand interactions

In the **Figure 2** we show the active sites for the PHF-tau protein and the respective sites where the box simulation was designed for the molecular docking study with the ligands.

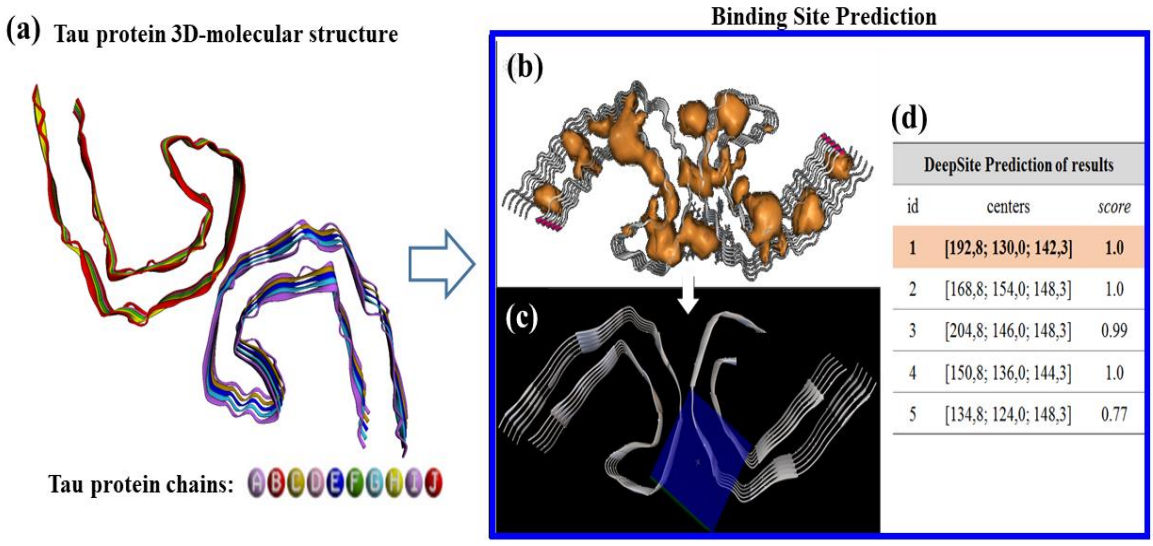


Figure 2. (a) Representation of three-dimensional structure of tau protein with the corresponding chains. (b) Representation of the predicted binding sites for tau protein. (c) Representation of the docking box simulation set for the most probable binding site of tau protein according with the maximum score in the DeepSite. (d) Table of results of the coordinates value obtained for the grid-box centers for the predicted binding sites highlighting the best ranked-binding site ID: 1. Besides, a ranking-based on the volume (V) in $Å^3$ for each predicted tau-protein binding site is reported like: site ID 1: V = 827.02 $Å^3$, site ID 2: V = 465.15 $Å^3$, site ID 3: V = 464.15 $Å^3$, site ID 4: V = 359.15 $Å^3$, site ID 5: V = 329.95 $Å^3$. In order to establish the relevance of each binding site of tau protein. Due to is well-know that the pocket cavity volume has a linear-correlation with the binding-site surface area, maximum intrinsic affinity of the ligand and its "druggability" or ability to strongly interact with a given binding site. More structural details can be found in supplementary materials as **Figure S1.** Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

As can be seen in **Figure 2b**, the receptor sites for PHF-tau are distributed throughout the tau macromolecule extent. However, for the prediction of the fit with the ligands was chosen a region among the best score according to DeepSite. The Cartesian coordinates for the grid-box center of simulation for the active site to evaluate the interaction between the ligands and

the tau protein were defined X = 192.8 Å, Y = 130.0 Å, Z = 142.3 Å as showed in **Figure 2** (**c** and **d**). Next, the box-simulation grid box-size were defined (X = 27 Å, Y = 27 Å, Z = 27 Å). Afterward, the molecular docking approaches to evaluate the binding interaction of the tau protein with the different ligands was carried out. The **Table 1** shows the obtained free energy of binding or affinity (kcal/mol) values for the best docking complexes with the corresponding root-means-square-deviations (R.M.S.D < 2 Å).

Table 1. Obtained free energy of binding (affinity; kcal/mol) and R.M.S.D (Å) for the best docking poses obtained of each PHF-ligand docking system evaluated and highlighting the best ranked docking system based on the free energy of binding (5: PHF – ^{18}F-THK5105+fullerol) denoted by the label symbol (*). Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

ID	Docking systems	inity (kcal/mol)	ISD (Å)
1	PHF – ^{18}F-FDDNP+fullerol	-5.9	1.356
2	PHF – ^{18}F-T807+fullerol	-6.0	1.617
3	PHF – ^{18}F-T808+fullerol	-5.7	1.344
4	PHF – ^{18}F-THK523+fullerol	-5.6	1.598
5*	**PHF – ^{18}F-THK5105+fullerol**	**-7.0**	**1.491**
6	PHF – ^{18}F-THK5117+fullerol	-6.8	1.628
7	PHF – ^{18}F-THK5351+fullerol	-6.6	1.360
8	PHF – ^{11}C-*N*-methyl lansoprazole+fullerol	-6.6	1.394
9	PHF – ^{11}C-PBB3+fullerol	-6.3	1.050

Next, we carried out the identification of relevant interactions between the docking systems evaluated and the relevant target residues involved as showed in the **Figure 3**.

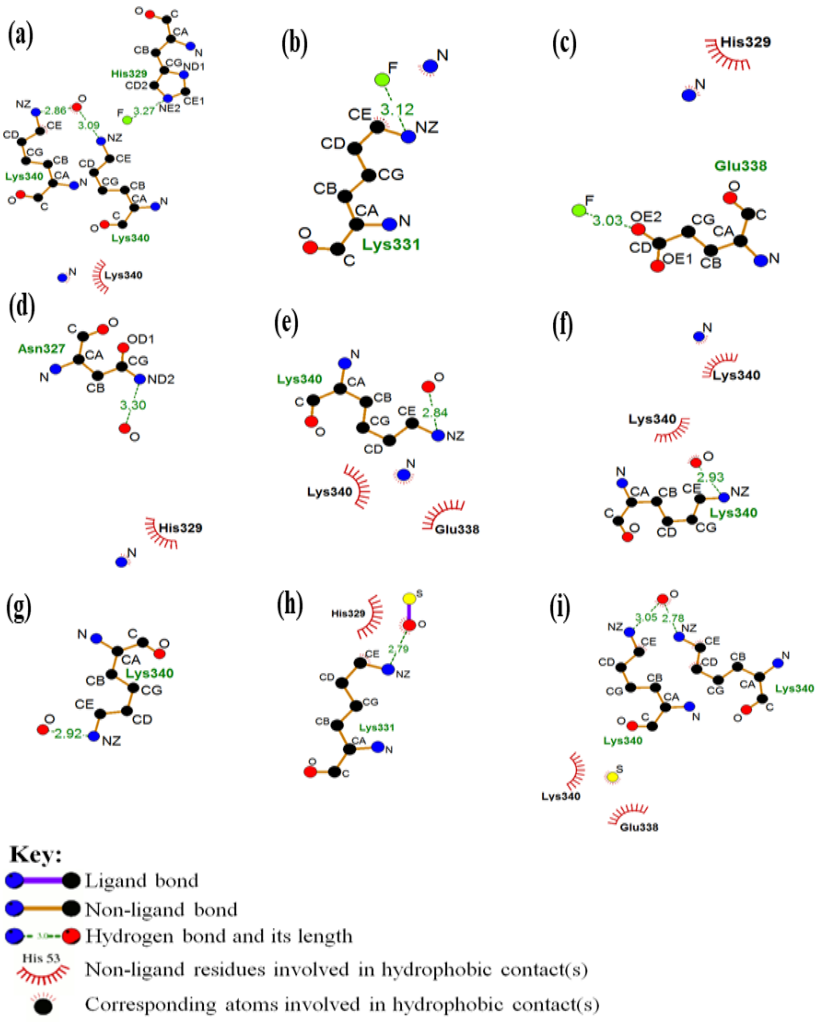


Figure 3. Representation of the most relevant docking interactions obtained for the best docking complexes with the corresponding key target residues of the tau protein in each evaluated docking system. Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

9.4 General Discussion

As observed in **Figure 3**, in system **(a)** the atoms of the ligand involved interactions of fluorine and oxygen atoms. For this instance, the fluorine interacted with the histidine residue of the tau protein through a hydrogen bond, with a distance of 3.3 Å. Yet, the oxygen interacted through hydrogen bonds with the lysine residues, with distances of 2.9 and 3.1 Å. In addition, was observed a hydrophobic interaction between the nitrogen atom and the lysine residue of the protein.

For system **(b)**, the region where direct interaction occurs involves the fluorine atom and the lysine residue of the protein, interrelating from a hydrogen bond. The distance from this

interaction was 3.1Å. Besides to hydrogen bonding, a hydrophobic interaction occurred between the nitrogen atom of the ligand and the carbon atom of the protein residue.

In the arrangement of the system **(c)**, it occurs a hydrogen bond of 3.0 Å of distance between the fluorine and the glutamate residue of the protein. In addition, a hydrophobic contact occurred between the nitrogen of the ligand and the Histidine residue of the macromolecule.

In the complex **(d)**, there was a hydrogen bond of 3.3 Å between the oxygen of the ligand and the nitrogen atom of the asparagine residue. There was also a hydrophobic contact between the ligand nitrogen and the histidine residue of the protein.

For best-ranked docking system (i.e., **(e)**: PHF – ^{18}F-THK5105+fullerol), we show that the oxygen atom of the ligand interacted through a hydrogen bond of 2.8 Å with the lysine residue. There were also two hydrophobic contacts between the nitrogen atom of the ligand and the lysine and glutamate residues of the macromolecule. Details are depicted in the **Figure 4** showing three-dimensional image representation of the whole docking system as PHF – ^{18}F-THK5105+fullerol represented in **(e)**.

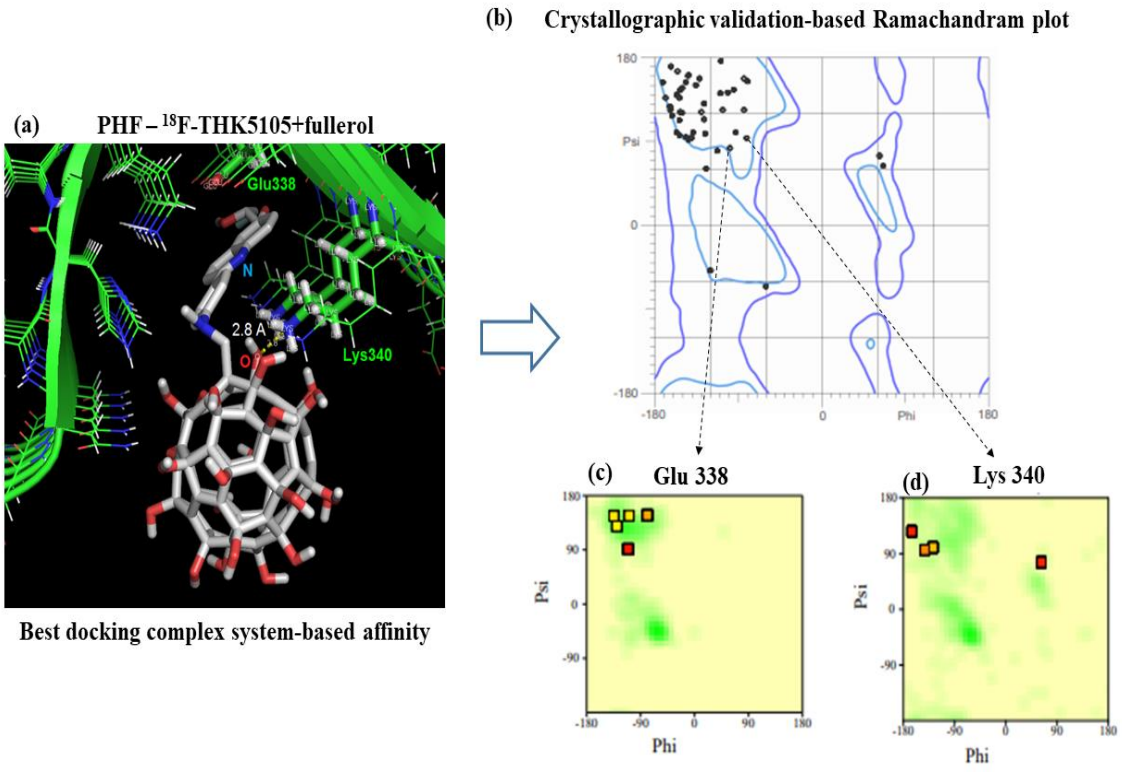


Figure 4. (a) Representation of the docking system (e) like a three-dimensional tau protein (PDB ID: 5O3L) binding site target residues (Glu338 and Lys340) interacting with atoms from the PHF – ^{18}F-THK5105 + fullerol. (b) Representation of Ramachandran diagrams obtained for all the residues composing the tau protein. Besides, the two arrows (c) and (d)

are depicted to show the individual Ramachandram driagram of the interacting target residues (Glu338 and Lys 340). All the possible combinations of torsion dihedral angles Phi (φ) *vs.* Psi (ϕ) are shown. Note the total absence of Ramachandran outlier residues for the three conditions evaluated (general and individual Ramachandran plot), usually unfavoured residues or torsion angles (Phi (φ) *vs.* Psi (ϕ)) are located outside the Ramachandran colored purple contour or green zone in the case of individual Ramachandran plots, if any. Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

As showed in the **Figure 4**, the crystallographic validation results show that the 100 % of the residues composing the tau protein can be categorized like conformationally favored residues with total absence of restricted flexibility. This fact presents a paramount relevance because prevents the presence of potential false positives results in the modeling docking interactions, as well as ensuring the quality of the results from the conformational point of view for the identified interactions with the target residues. Allowing to obtain a more proper evaluation on the conformational integrity of the target-residues potentially involved in the molecular recognition process toward biomedical detection purposes.

The remaining systems as **(f)**, there was a hydrogen bond of 2.9 Å between the oxygen of the ligand and the nitrogen of the lysine residue. Furthermore, two hydrophobic contacts occurred, one between the same oxygen atom of the ligand and the lysine residue, and another between the nitrogen atoms of the ligand the lysine residue. In the docking complex **(g)**, the main interaction of the system occurred through a hydrogen bond of 2.9 Å between the oxygen of the ligand and the nitrogen of the residue lysine.

In the arrangement **(h)** a hydrogen bond of 2.8 Å occurred between the oxygen of the ligand and the nitrogen of the lysine residue. In addition, there was a hydrophobic contact between the histidine residue and the same oxygen atom of the ligand.

Finally, in the system **(i)** it can be observed that two hydrogen bonds occurred with the oxygen atom of the ligand and lysine residues, with distances of 2.8 and 3.1 Å. In addition, there were also two hydrophobic contacts with the sulfur atom of the ligand and the residues lysine and glutamate.

In the analysis, the interactions that involved the systems were those of hydrophobic contact and hydrogen bonding. hydrophobic interactions occur between non-polar groups, and although are considered relatively weak interactions when compared to the others, these interactions are of great relevance in the complexation process between the protein and the

ligand (GUIMARÃES, 2012). In addition, all distances were less than 7 Å, indicating a higher likelihood of interaction in terms of affinity and less chance of false contacts (DURRUTHY et al., 2017).

The best complementarity feature according to docking criteria was obtained for the PHF – ^{18}F-THK5105+fullerol ligand system. As observed in the **Figure 4**, system **(e)**, the interface of this docking system is composed of the lysine residue (Lys340) interacting through a hydrogen bond of 2.8 Å with the oxygen atom of the PHF – ^{18}F-THK5105+fullerol ligand system. Based on this, the structural and electronic properties of the interaction interface of this complex were studied, as shown in **Figure 5**.

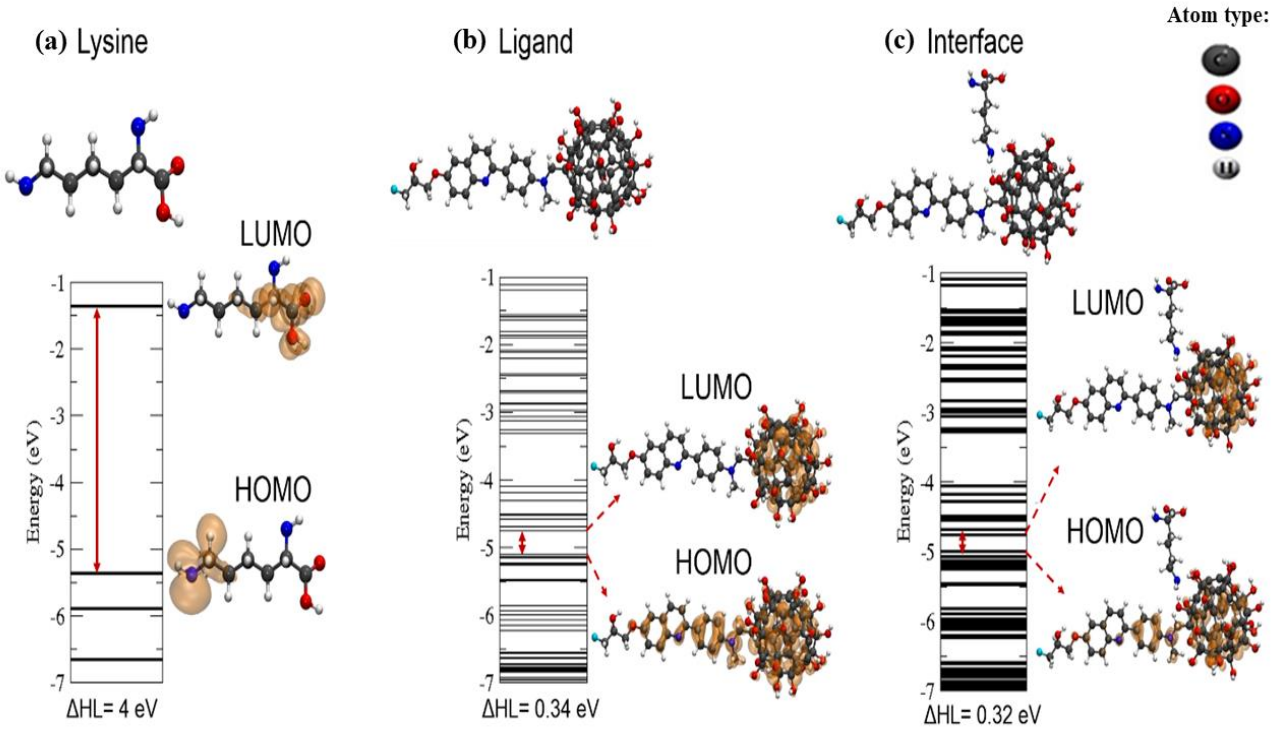


Figure 5. Representation of the interface of interaction for the best-ranked docking complex formed by binding site target residues (lysine; Lys340) interacting with atoms from the PHF – ^{18}F-THK5105 + fullerol. Herein, was represented the Lys340 residue because showed the high number of contact with the PHF – ^{18}F-THK5105 + fullerol ligand in the previous 2D-lig-plot analysis compared with the amino acid residue Glu338 in the tau protein. Then are represented the results of *ab initio* DFT simulation as: (a) energy levels, difference between the HOMO and the LUMO (ΔHL) and charges density of the isolated lysine molecule; (b) energy levels, ΔHL and charges density of the isolated ligand molecule; and (c) energy levels, ΔHL and charges density of the system interface. The value of the isosurface plot for LUMO

and for HOMO was 0.001 e-/Bohr³. Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

Table 2. Values of the electronic and structural parameters obtained for the interface of interaction for the best-ranked docking complex formed by binding site target residues (Lys340) interacting with atoms from the PHF – ^{18}F-THK5105 + fullerol. Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

Parameters	Value
Final distance	2.80 Å
Initial distance	2.83 Å
Binding energy	-0.9 eV

As showed above in the **Figure 5**, the electron density charges from LUMO and HOMO levels of the isolated structures (tau amino acid Lys34 and the PHF – ^{18}F-THK5105 + fullerol ligand) fit well with the docking results obtained for the interface of interaction **Figure 4(a)**. For this instance, the electron density charges were mainly concentrated over the ligand (PHF – ^{18}F-THK5105 + fullerol system) which is a relevant theoretical evidence for potential biomedical application toward the rational design of new diagnostic nanomarker. In addition, considering the results obtained in the **Table 2** we can observe that there were no significant changes between the final distance and the initial distance of the system, proving the results of the docking. Besides, the negative interaction energy (-0.9 eV) clearly indicates attraction between the structures (i.e., Lys340 tau-protein and ligand PHF – ^{18}F-THK5105 + fullerol). Lastly, the obtained theoretical results stronlgy suggest that the docking system is formed through a physical adsorption regime, which is an attribute of great relevance to be considered in the in the development of molecular detection system of high selectivity and specificity.

In addition, it is important to highlight that, theoretically, from the therapeutic point of view, the main advantages on using nanothecnology-based diagnosis markers over traditional methods of biomarker diagnosis we can refer the following: i) the possibility of significantly reducing the dose to be administered during the diagnostic procedures compared with conventional diagnostic procedures with better cost/benefit relationships, ii) significant increase of the bioavailability, as well as, optimal pharmacokinetics atributes namely: absorption, distribution, metabolism and elimination (ADME), iii) overcome the blood-brain barrier (BBB), which significantly limits the entry of micro-scale diagnostic compounds

compared with the nano-scale-based drugs allowing the more easy access to specific areas of the nervous tissue, in early stages of the Alzheimer's disease, iv) in consonance, the use of nanomarkers-based fullerol has theoretically an optimal balance between the lipophilic/hydrophilic properties over the conventional diagnostic systems based on its expected high lipid/water partition coefficient (Log P) that could favor the easy passage across the BBB, easy metabolism and elimination by the biochemical mechanisms of biotransformation (phase II) in the brain.

Lastly, the use of nanothecnology-based markers over traditional methods of diagnosis could significantly improve new therapeutic strategies of precision nanomedicine

References

1. AFSAHI, S. et al. Novel graphene-based biosensor for early detection of Zika virus infection. Biosensors and Bioelectronics, v. 100, p. 85-88, 2018.
2. BAKRY, R. et al. Medicinal applications of fullerenes. International Journal of Nanomedicina, v. 2, n. 4, p. 639-649, 2007.
3. BENADIBA, M. et al. Novos alvos moleculares para tomografia por emissão de pósitrons (PET) e tomografia computadorizada por emissão de fóton único (SPECT) em doenças neurodegenerativas. Revista Brasileira de Psiquiatria, v. 34, supl. 2, p. s125-s148, Out. 2012.
4. BENNETT, P.; LEIFER, M. D. Early Diagnosis of Alzheimer's Disease: Clinical and Economic Benefits. Journal of the American Geriatrics Society, v. 51, p. 281-288, May 2003.
5. BERMAN, H. M. et al. The Protein Data Bank. Nucleic Acids Research, v. 28, n. 1, p. 325-242, 2000.
6. BRUNDEN, K. R; TROJANOWSKI, J. Q; LEE, V. M. Y. Advances in tau-focused drug discovery for Alzheimer’s disease and related tauopathies. Nature Reviews, v. 8, p. 783-793, Oct. 2009.
7. CEPERLEY, D. M.; ALDER, B. J. Ground state of the electron gas by a stochastic method. Physical Review Letters, v. 45, n. 7, p. 566, 1980.
8. CHIEN, D. T. et al. Early Clinical PET Imaging Results with the Novel PHF-Tau Radioligand [F18]-T808. Journal of Alzheimer’s Disease, v. 38, n. 1, p. 171-184, 2014.
9. CULLUM B. M.; VO-DINH, T. The development of optical nanosensors for biological measurements. Trends in Biotechnology, v. 18, n. 9, p. 388-393, Sep. 2000.
10. DA RÓZ, A. L. et al. Grandes Áreas da Nanociência. 1 ed. Rio de Janeiro: Elsevier, 2015. Vol. 2.

11. DUGAN, L. L. et al. Buckminsterfullerenol free radical scavengers reduce excitotoxic and apoptotic death of cultured cortical neurons. Neurobiology of Disease, v. 3, n. 2, p. 129-135, 1996.
12. DURRUTHY, M. G. et al. Decrypting Strong and Weak Single-Walled Carbon Nanotubes Interactions with Mitochondrial Voltage-Dependent Anion Channels Using Molecular Docking and Perturbation Theory. Nature Scientific Reports, v. 7, n. 1, p. 1-19, 2017.
13. ERKEKOGLU, P.; GIRAY, B. K.; BASARAN, N. 3R Principle and Alternative Toxicity Testing Methods. Fabad Jounal of Pharmaceutical Sciences, v. 36, p. 101-117, 2011.
14. GASSER, A. L.; SALAMIN, V.; ZUMBACH, S. Late life depression or prodromal Alzheimer's disease: Which tools for the differential diagnosis? Encephale, Apr. 2017.
15. GUIMARÃES, C. R. W. As múltiplas Contribuições para a Complexação Proteína-Ligante: Consequências em *Drug Design*. Revista Virtual de Química, v. 4, n. 4, p. 348-364, 2012.
16. HOHENBERG, P.; KOHN, W. Inhomogeneous electron gas. Physical review, v. 136, n. 3B, p. B864, 1964.
17. HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD - Visual Molecular Dynamics. Journal of Molecular Graphics, v. 14, n. 1, p. 33-38, 1996.
18. JACK, C. R. et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia, v. 7, n. 3, p. 257-262, May 2011.
19. JIMÉNEZ, J. et al. DeepSite: protein-binding site predictor using 3D-convulation neural networks. Bioinformatics, v. 33, n. 19, 2017.
20. LASKOWSKI, R. A.; SWINDELLS, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model, v. 51, n. 10, p. 2778-2786, Sep. 2011.
21. MASTERS, L. C. et al. Alzheimer's disease. Nature Reviews, v. 1, n. 15056, p. 1-18, Oct. 2015.
22. MATSUNAGA, A.; YONEDA, M. Dementia due to Endocrine Diseases. Brain and Nerve, v. 68, n. 4, p. 399-405, Abr. 2016.
23. PASSINI, E. et al. Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity. Frontiers in Physiology, v. 8, n. 668, p. 1-15, Sep. 2017.
24. PRINCE, M.; BRYCE, R.; FERRI, C. World Alzheimer Report 2011: The benefits of early diagnosis and intervention. London: Alzheimer's Disease International, 2011. Disponível em: <https://www.alz.co.uk/research/WorldAlzheimerReport2011.pdf>. Acesso em: 28 de maio de 2017.

25. PRINCE, M.; BRYCE, R.; FERRI, C. The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International, 2015. Disponível em: < http://www.worldalzreport2015.org/downloads/world-alzheimer-report-2015.pdf>. Acesso em: 28 de maio de 2017.
26. PRINCE, M. et al. Improving healthcare for people living with dementia: coverage, quality and costs now and In the future. London: Alzheimer's Disease International, 2016. Disponível em: < https://www.alz.co.uk/research/WorldAlzheimerReport2016.pdf>. Acesso em: 28 de maio de 2017.
27. SANTOS, L. J. et al. Fulereno [C60]: química e aplicações. Química Nova, v. 33, n. 3, p. 680-693, 2010.
28. SCHELTENS, P. et al. Alzheimer's disease. The Lancet, v. 388, n. 10043, p. 505-517, Feb. 2016.
29. SLOANE, P. D. et al. The public health impact of Alzheimer's disease, 2000-2050: potential implication of treatment advances. Annual review of public health, v. 23, n. 1, p. 213-31, 2002.
30. SOLER, J. M. et al. The SIESTA method for ab initio order-N materials simulation. Chemical Reviews, v. 14, p. 2745-2779, Mar. 2002.
31. TOUHAMI, A. Biosensors and Nanobiosensors: Design and Applications. Nanomedicine. Brownsville: [s.n.], p. 374-400, 2015.
32. TROTT, O; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comp Chem, v. 31, n. 2, p. 455-461, 2010.
33. VILELA NETO, O. P.; PACHECO, M. A. C. Nanotecnologia computacional inteligente: Concebendo a engenharia em nanotecnologia. 1 ed. Rio de Janeiro: Interciência, 2012.
34. VILLEMAGNE, V. L. et al. Tau imaging: early progress and future directions. Lancet Neurology, v. 14, p. 114-124, Jan. 2015.
35. WANG, Y. et al. Pubchem BioAssay: 2017 update. Nucleic Acids Research, v. 45, p. 955-963, Nov. 2016.
36. YANG, Y. et al. Aptamer-functionalized carbon nanomaterials electrochemical sensors for detecting cancer relevant biomolecules. Carbon, v, 129, p. 380-395, 2018.

Supplementary materials

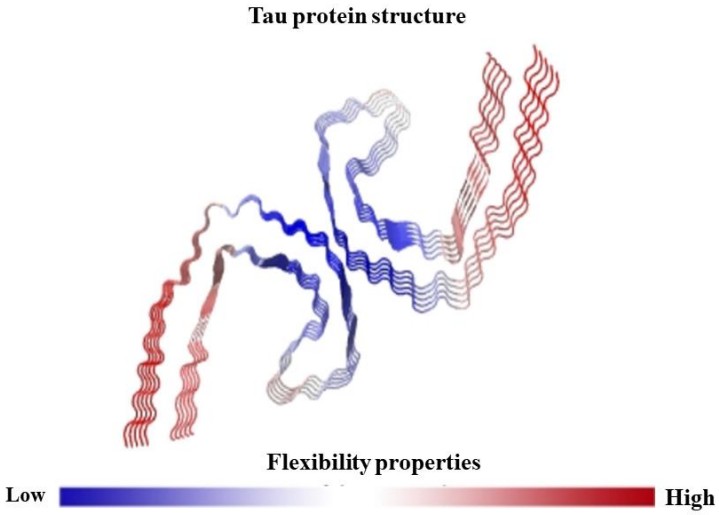


Figure S1. Schematic representation of the tau protein flexibility properties. On the bottom the colored bar is to represent the size of fluctuations of the tau protein residues showing from low-flexibility (blue) to high-flexibility (red). Note that the best ranked binding site (ID:1) presents intrinsic rigidity (blue) simulated in physiological conditions. Source MDPI article, please visit: https://doi.org/10.3390/biophysica1020007 (accessed on 24 March 2021)

Chapter 10.

Patrícia Viera de Oliveira*, Luiza Goulart, Cláudia Lange dos Santos, Jussane Rossato, Solange Binotto Fagan*, Ivana Zanella, M. Natália D. S. Cordeiro, Juan M. Ruso, **Michael González-Durruthy***. Computational Modeling of Environmental Co-exposure on Oil-Derived Hydrocarbon Overload by Using Substrate-Specific Transport Protein (*TodX*) with Graphene Nanostructures. Current Topics in Medicinal Chemistry, 20, 1-18. 2020. doi.org/10.2174/1568026620666200820145412.

Journal Impact Factor (JIF) (2020): 3.218.

CiteScore (2020): 5.8

Author Contribution M.G.-D.; Conceptualization, methodology-based on molecular docking,binding cavities predictions, evaluations on flexibility properties of TodX, Ramachandran crystallographic validation, 2D/3D-lig-plot analysis, modeling heterogeneous multicomponent docking with multiplicity states, DFT results interpretation, and writing—original draft preparation. *Corresponding author

Journal authorization:

LICENSED CONTENT

Publication Title	Current topics in medicinal chemistry	Rightsholder	EUREKA SCIENCE (FZC)
Article Title	Corrigendum to: Computational Modeling of Environmental Co-exposure on Oil-Derived Hydrocarbon Overload by Using Substrate-Specific Transport Protein (TodX) with Graphene Nanostructures	Publication Type	e-Journal
		Start Page	839
		Issue	9
		Volume	21
Date	01/01/2001		
Language	English		

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	13	Title of the article/chapter the portion is from	Corrigendum to: Computational Modeling of Environmental Co-exposure on Oil-Derived Hydrocarbon Overload by Using Substrate-Specific Transport Protein (TodX) with Graphene Nanostructures
Editor of portion(s)	de Oliveira, Patrícia Viera; Goulart, Luiza; dos Santos, Cláudia Lange; Rossato, Jussane; Fagan, Solange Binotto; Zanella, Ivana; Cordeiro, M. Natália D.S.; Ruso, Juan M.; González-Durruthy, Michael	Author of portion(s)	de Oliveira, Patrícia Viera; Goulart, Luiza; dos Santos, Cláudia Lange; Rossato, Jussane; Fagan, Solange Binotto; Zanella, Ivana; Cordeiro, M. Natália D.S.; Ruso, Juan M.; González-Durruthy, Michael
Volume of serial or monograph	21		
Page or page range of portion	839	Issue, if republishing an article from a serial	9
		Publication date of portion	2021-05-26

RIGHTSHOLDER TERMS AND CONDITIONS

If your permission request relates to Open Access content, published under the CC BY 4.0 license, you don't need to take permission from Bentham Science for reuse, as long as the original publication and Bentham Science are correctly credited.

Total Items: 1 — Total Due: 0,00 EUR

Accepted: All Publisher and CCC Terms and Conditions

Chapter 10.

Modeling Simultaneous Interactions of Oil-Derived Hydrocarbon with the Transport Protein (*TodX*) and Graphene Nanostructures

Abstract

Bioremediation strategies widely uses living organisms to remove persistent contaminants like oil-derived under oil spill conditions to the environment mainly soils and water. Here, a new mechanistic approaches by applying a combined methodology as Molecular Docking and Density Functional Theory were proposed to modeling the interactions of a heterogeneous mixture composed by oil-derived hydrocarbons, pristine and oxidized graphene nanostructures, and the substrate-specific transport protein (TodX) from *Pseudomonas putida* toward new bioremediation strategies. The theoretical evidences strongly suggest that the binding interactions are mainly based in non-covalent hydrophobic interactions characterizing a physical adsorption mechanism that mimics the "*Trojan-horse effect*". These results open new horizons to improve bioremediation strategies in overload-saturation conditions against oil-spills expanding the use of nanotechnologies in the context of environmental health and safety.

10.1 Introduction

One of the largest oil-spills in the human history was recorded in 2010 in the Gulf of Mexico on the Deepwater Horizon platform, where approximately 800 million liters of oil were spilled into the sea.[1] This environmental catastrophe caused severe impacts over the health of birds, fish, and marine mammals with a high-mortality rate in this ecosystem, including human health where the prevalence of neurological disorders and other environmental diseases currently persist.[2,3] In the present study, we tackle the environmental modeling based on the use of detoxifying microorganisms (*Pseudomonas putida*) which present efficient biochemical mechanisms to selectively-reduce the oil-derived hydrocarbons concentrations. The *Pseudomonas putida* is a bacterium which has been studied for many environmental researchers, due to its potential application in the bioremediation field. *Pseudomonas putida* is naturally adapted to actions against crude oil spillage in the environment.[4] Particularly, the outer membrane transport protein

(TodX) family from *Pseudomonas putida* includes several proteins that are involved in the efficient degradation of aromatic hydrocarbons.[5-7] The biochemical function of the TodX protein allows the hydrocarbons detoxification and passage through the cell membranes.

However, the TodX hydrophobic-transport activity could be compromised in overload conditions of several pollutants.[5-7] Previous studies performed by Belchik and colaborators[8], support that a similar protein-like TcpY facilitates the transport of polychlorophenols across the outer membrane of gram-negative bacterium,[8] and likewise COG4313 proteins are possible outer membrane transporters of hydrophobic aromatic compounds.[8] Besides, Van Den Berg[9] investigated the FadL family channels, as carrier of hydrophobic molecules that work according to a lateral diffusion mechanism that contribute to the detoxification of hydrophobic aromatic compounds.

By the other hand, interesting work performed by Hearn et al.[10] tackle the structural study of the two layers of the outer membranes using *Pseudomonas putida* (TodX) and *Ralstonia pickettii* (TbuX) to address the transport mechanism of hydrocarbons using *in vitro* assays, suggesting that hydrocarbons transport mechanism is mainly based on hydrophobic transport from the extracellular environment to the periplasm through a channel hatching domain of the TodX.[10] Besides, high-performance liquid chromatography measurements strongly suggest that this microorganism *Pseudomonas putida (*TodX*)* may be useful in the biodegradation of polycyclic aromatic hydrocarbons PAHs in petrochemicals.[10,11] Following this idea, carbon nanomaterials like graphene nanostructures (pristine-Gr and oxidized-Gr) could provide new alternatives for environmental bioremediation when combined with detoxifying microorganisms (TodX protein from *Pseudomonas putida*), due to the high surface reactive ability to adsorb aromatic hydrocarbons compounds.[11-14] The pristine-Gr structures present a significantly high- affinity by aromatic hydrocarbons compounds due to the occurrence of hydrophobic π-π interactions in the reactive surface of these nanomaterials. In addition, previous studies show that the graphene functionalization-based oxidation processes significantly increase the absorption properties. Because oxidized-surface moieties (epoxy, OH and COOH) present in the oxidized-Gr reactive surface present high-affinity to strongly interact with polycyclic aromatic hydrocarbons (PAHs).[11,12]

In this context, the graphene nanostructures, could provide a promising bioremediation alternative for the treatment of water contaminated by oil and its derivatives in overload conditions.[10-12] It is well-known that the specific properties of graphene nanostructures

can provide a high adsorption capacity initially coined with the term "*Trojan-horse effect*".[10-15] Several studies have recognized the *Trojan-horse-like effects* induced by carbon nanomaterials as an important mechanism to reduce lethal concentrations and negative impact of many organic pollutants on the environmental compartment (soil and water).[15] Particularly, graphene nanostructures (pristine-Gr and oxidized-Gr) present a high adsorption ability, which is intentionally employed in several applications involving molecular recognition, nanomedicine, and bioremediation.[14-18]

Theoretical studies have been proposed, in this work, based on environmental modeling and simulations, that can qualitatively reproduce the complex processes that occur in the environmental compartments using mathematical and physical models to describe the dynamic of the systems and to predict their behavior. At present there are no studies in the literature addressing the analysis of protein-ligand and ligand-ligand interactions on detoxifying proteins and multiple-ligand-like (environmental pollutants) with a heterogeneous environment. For this purpose, the present study propose, for the first time, a new computational approach on the outer membrane protein transport (TodX) from the detoxifying microorganism (*Pseudomonas putida*) interacting with oil-derived hydrocarbons (n-butylbenzene, n-octane and methylnaphthalene) and graphene nanostructures (pristine-Gr and oxidized-Gr) to mimicking overload conditions, like oil-spill by using a new multicomponent docking approach with *ab initio* Density Functional Theory (DFT), in order to tackle the theoretical study on bioremediation strategies.[14-19] In this regard, the present study propose new relevant information underlying the mechanisms of environmental co-exposure, and, at the same time, open new horizons for the nanotechnology applications in the context of environmental modeling and health safety.

10.2 Materials and Methods

10.2.1 TodX actives binding-sites prediction

The potential TodX binding-sites were predicted through DeepSite[20], which identify TodX-cavities as relevant catalytic sites that is likely to bind to a small ligand into the Van der Waals surfaces. To this end, DeepSite considers all the molecular descriptor related to the proteins TodX using machine-learning algorithm based on 3D-deep convolutional neural networks (DCNNs).[20] The DeepSite method is a robust and validated deep learning library with an extensive test set based on > 6500 proteins of the scPDB database which allows an unequivocal identification of the catalytic binding site in any

protein. In this regard, the binding pocket predictions, as well as, the TodX mesh volumetric-map obtained from DeepSite was used to establish the Cartesian coordinates of the docking box simulation like TodX grid box size, with X = 22 Å, Y = 18 Å, Z = 32Å dimensions and with TodX-box simulation center as X = 29.615 Å, Y = 63.638 Å and Z = 28.934 Å, with option of exhaustiveness of 100 (average precision).[21, 22]

10.2.2 Ramachandran plot validation

To validate the 3D x-ray crystallographic structure of the transport protein *Pseudomonas putida* (TodX) pdb.model the Ramachandran plot as a general case and the corresponding quality assessment was performed.[23] In order to prevent obtaining false positives on flexible-docking complexes. This methodology allows to verify the absence of restricted-flexibility for each residue from the TodX protein based on the identification of allowed and disallowed torsion values of Ψ (*Psi*) versus Φ (*Phi*).

10.2.3 Performing elastic network models on TodX binding sites

This computational algorithm evaluates the signal communication efficiency between two or more binding active site of a given protein (TodX) by describing the residues network communication like Hookean potential (or 'springs') based on elastic normal mode (ENM) analysis of the interaction potential (*U*) obtained from the different sites.[24-26] The mathematical framework of this approach will be discussed in details in the next section.

10.2.4 Molecular docking simulation with multicomponent binding systems

In order to study the binding-interactions between the ligands (*i.e.*., oil-derived hydrocarbons: n-butylbenzene, n-octane and methylnaphthalene and the graphene nanostructures: pristine-Gr and oxidized-Gr) with the outer membrane toluene transporter (TodX) from *Pseudomonas putida.* The 3D-crytallographic structure file like TodX protein was retrieved from the RCSB Protein Data Bank (PDB) like PDB ID: 3BRZ with resolution of 3.2 Å.[27] Before the docking approach, the TodX-protein was optimized by using the AutoDock Tools 4 software for AutoDock Vina.[27-29]

Next, the TodX-hydrogen atoms were added according to appropriate hybridization geometry by adding the Gasteiger partial atomic charges, protonation states followed by bond orders assignment and rotatable bonds of the TodX protein.pdb file.[27-30]

Afterward, the ligand structures like methylnaphthalene (PubChem CID: 7002), n-butylbenzene (PubChem CID: 7705), n-octane (PubChem CID: 356), [32] and the graphene

nanostructures (pristine-Gr and oxidized-Gr) like the mol2. input files were optimized by using the MOPAC extension for geometry optimization based on the AM1-Hamiltonian method.[31-32] Herein, it is important to clarify that the toluene ligand even though is the most recognized substrate of the toluene transport protein *Pseudomonas putida* (TodX) we did not take it into account for the purposes of this study, focusing our attention in the hydrophobic oil-derived hydrocarbon TodX substrates with greater degree of environmental persistence and bioaccumulation like methylnaphthalene, n-butylbenzene, and n-octane.[32]

According to the main objective of this work to evaluate the interactions mechanism from the formed complexes, the free energy of binding or affinity (FEB, kcal/mol) was obtained considering the following systems: **i**) TodX + n-butylbenzene, ***ii***) TodX + n-octane, ***iii***) TodX + methylnaphthalene, ***iv***) TodX + Gr-pristine and ***v***) TodX + oxidized-Gr. Furthermore, to mimicking the environmental co-exposure with the oil-derived hydrocarbons with two graphene nanostructures the following two heterogeneous-multicomponent docking systems were considered: ***vi***) TodX + (methylnaphthalene/n-butylbenzene/n-octane) + Gr-pristine and ***vii***) TodX + (methylnaphthalene/n-butylbenzene/n-octane) + oxidized-Gr. For this instance, the Autodock Vina scoring function was implemented.[27] Herein, the implemented the Vina scoring function combines the knowledge-based potential and empirical information from experimentally-validated binding affinity measurements with default Amber force-field thermodynamic parameters.[27-29] For this instance, the Vina scoring function considers optimal-linear Gibbs free energy of binding affinity ($\Delta G \approx$ FEB dock). It is important to note that, overall docking force field parameters are based on distance-dependent atom-pair interactions (d_{ij}). The free energy of binding (FEB kcal/mol as output of docking results) was defined by ΔG_{bind} values for all docked poses (TodX-ligands) like a sum of the individual molecular mechanics terms of standard-chemical potentials as: Van der Waals interactions (ΔG_{vdW}), hydrogen bond ($\Delta G_{H\text{-}bond}$), electrostatic interactions ($\Delta G_{electrost}$), and intra-molecular ligands interactions ($\Delta G_{internal}$).[27-29] See the general thermodynamic **equations (1)** and **(2)**.

$$FEB_{dock} \approx \Delta G_{bind} = \Delta G_{vdW} + \Delta G_{H-bond} + \Delta G_{eletrost} + \Delta G_{internal} \quad \textbf{(1)}$$

$$FEB_{dock} \approx \Delta G_{bind} = \Delta G_{VdW} \sum_{lig-TodX} \left(\frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^{6}}\right) + \Delta G_{H-bond} \sum_{lig-TodX} E(t) \left(\frac{C_{ij}}{d_{ij}^{12}} - \frac{D_{ij}}{d_{ij}^{10}}\right) + \Delta G_{elec} \sum_{lig-TodX} 332.0 \frac{q_i q_j}{\epsilon(d_{ij}) d_{ij}} + \Delta G_{int} \left\{\left\{\sum_{lig} \frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^{6}} + \sum_{lig} E(t) \times \left(\frac{C_{ij}}{d_{ij}^{12}} - \frac{D_{ij}}{d_{ij}^{10}}\right) + \sum_{lig} 332.0 \frac{q_i q_j}{4 d_{ij} d_{ij}} + \sum_{lig} \gamma_k (1 + \cos(\omega_k \theta_k - \theta_{0k}))\right\}\right. \quad \textbf{(2)}$$

The affinity (ΔG_{bind}) values obtained from the docking complexes were categorized-like energetically-unfavorable when the $\Delta G_{bind} \geq 0$ kcal/mol, indicating an either extremely low or complete absence of affinity of the oil-derived hydrocarbons and/or graphene nanostructure interaction complexes; otherwise are categorized-like medium to high-affinity of binding. Additional details on the thermodynamic force field parameters used in this study can be consulted in Autodock Vina scoring function.[27-39]

By the other hand, the ligand conformations with the lowest Gibbs free energies of binding (*i.e.*, the more negatives FEB values) were retrieved like outputs. The best root-mean-square deviation (*R.M.S.D*) was considered as a criterion of correct docking pose accuracy below 2Å according to the **equation (3)**.

$$RMSD(pose_{i-lig}, pose_{j-TodX}) = \sqrt{\frac{\sum n\left((atom_{(i-lig)}) - (atom_{(j-TodX)})\right)^2}{n}} \quad \textbf{(3)}$$

10.2.5 DFT simulation

The *ab initio* DFT methodology, implemented with the SIESTA code (Spanish Initiative for Electronic Simulations with Thousands of Atoms)[41], was used combined with docking simulation experiments. In this regard, two DFT-simulation groups presenting the most stable binding configuration three DFT-interacting systems each one were considered like: **Group #1**: ***i***) pristine-Gr + n-butylbenzene, ***ii***) pristine-Gr + n-octane, ***iii***) pristine-Gr + methylnaphthalene and the **Group #2**: ***i***) oxidized-Gr + n-butylbenzene, ***ii***) oxidized-Gr + n-octane, ***iii***) oxidized-Gr + methylnaphthalene considering the optimized systems obtained from the most stable final configuration. For this instance, DFT-algorithm performs self-consistent calculations to solve the Kohn-Sham equation and evaluates the resulting properties of the interactions studied.[42] For this instance, double-base function polarized (DZP) was used for the exchange-correlation potential using a Local Density

Approximation (LDA) parameterized by Perdew and Zunger[43] with cutting radius of 300 Ry for the mesh (grid) of the real space interaction. The atomic structures were relaxed until the residual forces were less than 0.05 eV / Å for all the atoms of the system. The interactions between the valence electrons and the ionic core was treated within the pseudopotential method as proposed by Troullier and Martins.[44] For each DFT-simulation group (**Group #1 and Group #2**), several DFT-parameters were evaluated namely: the interaction energy of the ligands like (E_b), smallest interatomic distance (d), and the charge transfer (ΔQ). Furthermore, the electronic band structures, as well as, the associated conduction band minimum (CBM) and the valence band maximum (VBM), with respect to the Fermi level were determined in the cases of graphene nanostructures (pristine-Gr and oxidized-Gr). By the other hand, in the case of the oil-derived-hydrocarbons (methylnaphthalene, n-butylbenzene or n-octane) the associated energy difference between the highest occupied molecular orbital and the lowest unoccupied molecular orbital (HOMO-LUMO gap energy) was obtained. For this instance, the binding energy (E_b) of the cited DFT-interacting systems were calculated according to the following **equation 4**:

$$E_b = -\{E(A+B) - E(A) - E(B)\} \quad (\mathbf{4})$$

Where, the term E(A+B) is the total energy of the interaction between an evaluated hydrocarbon (methylnaphthalene, n-butylbenzene or n-octane) with a graphene nanostructure (pristine-Gr or oxidized-Gr), the term E(A), represents the total energy of an isolated hydrocarbon molecule, and the last E(B) corresponds to the total energy of the isolated graphene nanostructure.

10.3 Results and discussion

Bacterial biodegradation of oil-derived hydrocarbons is a novel and important process for environmental remediation, which requires the hydrophobic transport of these compounds across the bacterial cell membrane. In this regard, we report for the first time a computational study toward to improve the current knowledge about environmental detoxification mechanisms using *Pseudomonas putida*.[6-10] Specifically, the modeling of interactions of the outer membrane transport protein TodX, which have been implicated in the passage of hydrophobic molecules and biodegradation of oil-derived hydrocarbons from the extracellular environment towards the periplasm.[6-10] In this regard, we carried

out a theoretical modeling by adapting the traditional molecular docking procedures by performing a new heterogeneous multicomponent docking approach in accordance with our objective.[33,34] Then, we assume that the canonical ligand-protein (1:1) proportion (*i.e.*, one hydrocarbons molecule: one protein TodX) fit with under-exposure conditions with easy biodegradation by TodX.[22,28,29]

On the other hand, we considered the modeling based on the heterogeneous mixture of oil-derived hydrocarbons like methylnaphthalene, n-butylbenzene and n-octane in the multicomponent ligand-protein evaluation. Mimicking environmental overload conditions (high hydrocarbon concentrations). [33,34] In this context, we introduce graphene nanostructures to tackle the docking and DFT study to model potential physico-chemical adsorption of the cited oil-derived hydrocarbons concomitantly interacting with the transport protein TodX in equilibrium conditions.[33,34]

To start the study, one of the most important steps to ensure quality of modeling data consists in the prediction/identification of the suitable TodX binding-sites along with an appropriate crystallographic-structural validation and flexibility properties of this protein. In this concern, in the present work the TodX binding-sites predictions were carried out by applying *ezPocket* package which use Voronoi tessellation method to detect tiny cavities and collecting information-based crystallographic descriptors about the pocket topology with high probability to bind small ligands.[20,21] The binding cavity detection method (*fpocket*) generates cartesian coordinates-like volumetric maps that allows setting the docking box-simulation on the TodX binding site. For this instances to binding sites were identified like: TodX_Site 1 grid box-size with dimensions of X= 54 Å, Y= 54 Å, Z= 54 Å and TodX grid box-center X= 31.08 Å, Y= 23.18 Å, Z = 51.41 Å with volume of 769.92 $Å^3$ and the TodX_Site 2 grid box-size with dimensions of X= 64 Å, Y= 64 Å, Z= 64 Å and TodX grid box-center X= 26.43 Å, Y= 47.63 Å, Z = 27.09 Å with volume of 3846.61 $Å^3$. [20,21] Furthermore, the flexibility properties and collective motions-like anisotropic fluctuations were evaluated from the whole TodX structure.[24-26] See **Figure 1**.

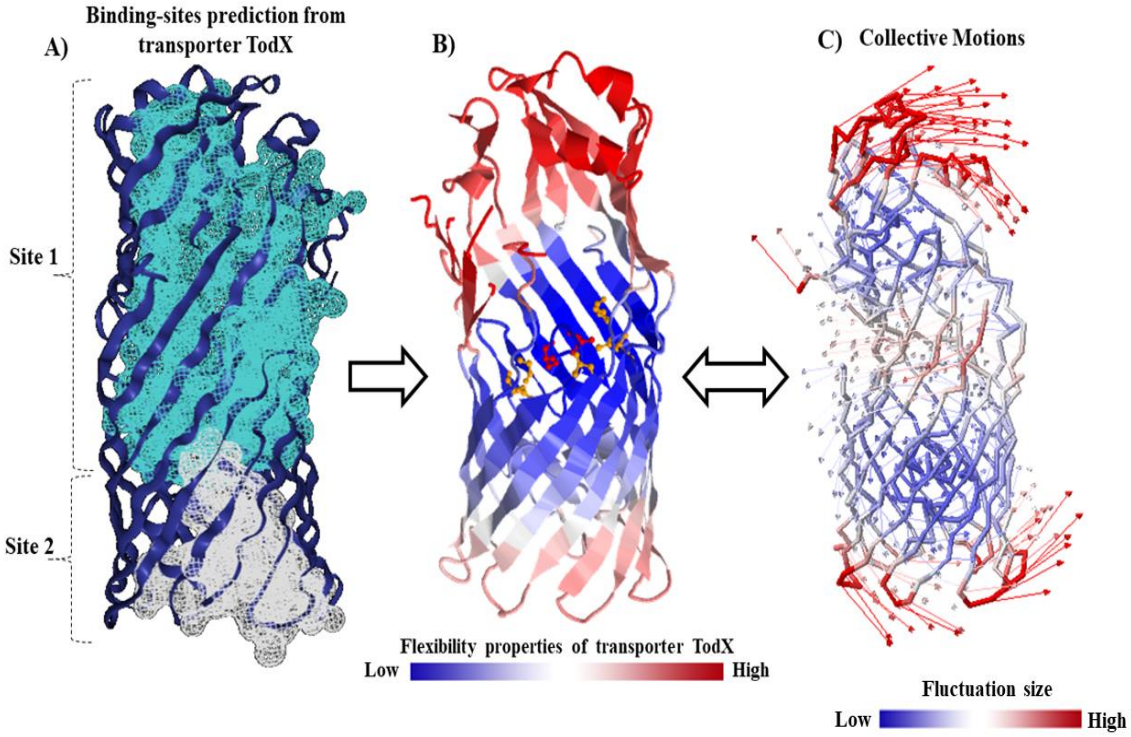


Figure 1. **A)** Representation of ezPocket calculation/prediction of topological-cavities like embedded-membrane binding site 1 (light blue) and extracellular domain also named hatch domain or site 2 (white) from the outer membrane transport protein (TodX) family from Pseudomonas putida represented as volumetric Van der Waals surface/mesh representation. **B)** Representation of TodX flexibility properties like 3D-colored structure based on the size of fluctuations-driven according to the slowest vibration-modes from TodX-residues from low-flexibility (blue) to high-flexibility (red). Herein, the catalytic binding residues are in the region of the site 1 in the middle of the channel labeled-yellow and red. **C)** On the far right, atomic positional fluctuations illustrating the correlated motions between the two previously identified binding sites like site 1 and site 2. Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

Afterward, to validate the 3D-X-ray crystallographic structure of the TodX-structure the Ramachandran plot diagram coupled to quality assessment was performed to avoid false positives in the flexible docking results.[23] See **Figure 2**.

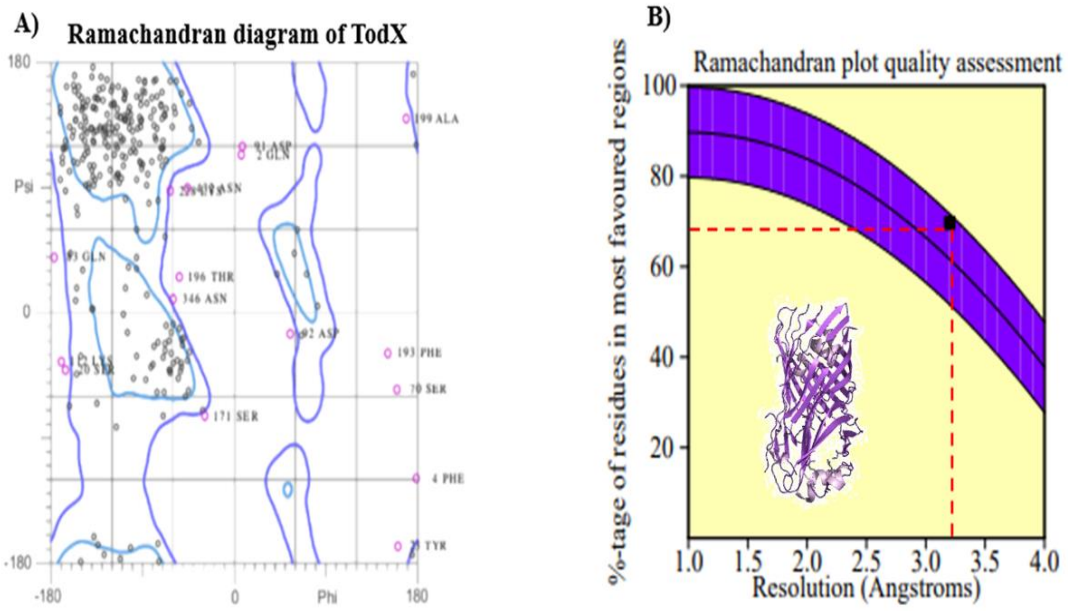


Figure 2. **A)** Representation of Ramachandran analysis with the spatial distribution of Ramachandran outliers (pink balls) in the modeled pdb x-ray structure of TodX protein. All the possible combinations of dihedral angles of torsion like Psi vs. Phi of each amino acid residue of TodX are showed. **B)** Results of Ramachandran quality assessment (3brz.pdb model quality) measured by the percentage of the TodX-residues which are in the most favored, or core, regions of the Ramachandran plot vs. resolution (Å). For this instance, the TodX crystallographic structure exhibits an acceptable crystallographic quality with more than 70% of the residues placed in the favored region including the site 1 and site 2 (red-dotted line). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

Ramachandran evaluation is a 2D-projection on the plane from the 3D-crystallographic TodX model.pdb where all the possible TodX binding sites conformations of each residue are defined into the plot according to the torsion dihedral angles Psi *vs.* Phi around the TodX peptide-bond residues.[23] In this context, allowed torsion values of the cited dihedral angles are placed within the Ramachandran colored purple contour (as conformational-favored TodX-residues).[23] Otherwise, are considered as TodX-sterically-disallowed residue associated with torsion values of dihedral angles Psi *vs.* Phi which appear outside of the Ramachandran colored purple contour (like conformational non-favored amino acid residues). Following the sequence of modeling interactions, we performed the *in*

silico experiment in order to get the free energy of binding (FEB, kcal/mol) for the complexes formed between the TodX protein and the oil-derived hydrocarbons, mimicking low or high-concentration in the absence or presence of graphene nanostructure (pristine-Gr and oxidized-Gr).

Herein, it is important to refer some environmental issues that surrounding environmental processes before moving directly to models these complex interactions. In this sense, for the best of our knowledge, there is no evidences considering the environmental modeling using graphene nanostructure with detoxifying microorganisms. Although there are very interesting works based on experimental nanoecotoxicological evaluations, which have been able to reproduce with a certain level of precision the *Trojan-horse phenomenon* induced by carbon nanomaterials in co-exposure conditions.[15]

Firstly, was performed the flexible docking interaction mimicking low hydrocarbons concentrations.[29-34] Hydrocarbons that derived from petroleum such as oil, gasoline, and diesel-fuel are among the most commonly occurring and widely-used contaminants in the environment.[45] Crude oil occurs inside the earth and is composed by a complex mixture of natural compounds composed mostly of hydrocarbons containing only H-atom and C-atoms. The remaining elements of sulfur, nitrogen, and oxygen are less than 3% of the petroleum.[45,46] Petroleum oil-derived products are released within the environment accidentally and intentionally at all steps of its use like exploration, production, transportation, storage, use and disposal. The most volatile subcomponents of petroleum are biodistributed preferentially into the atmosphere and hydrosphere inducing the well-known bioaccumulation and bioconcentration processes which exceed the biotransformation capacity of the of the environmental compartments mainly water or can impact the trophic chains (including the Homo sapiens trophic chain) by inducing the biomagnification process in the most of cases.[47-49] This event can be defined like the increase of the concentration of pollutant (oil-derived hydrocarbons) through the different levels of the trophic chain. [47-49] Then, after defining this ecotoxicological concepts we can pass to explain our computational approaches.

It is important to note that the low-concentration condition was modeled according to the theoretical protein: ligand proportion (1:1) (*ie.,* one hydrocarbons molecule: one protein TodX) mimicking an easy biodegradation by TodX (*ie.,* docking interaction with the previously predicted catalytic binding residues in the TodX site 1) not involving potential hydrocarbon-hydrocarbon interactions as usually happens in mixtures released into the

environment in high-concentration. The results of the flexible docking procedures applied are depicted in the **Figure 3**

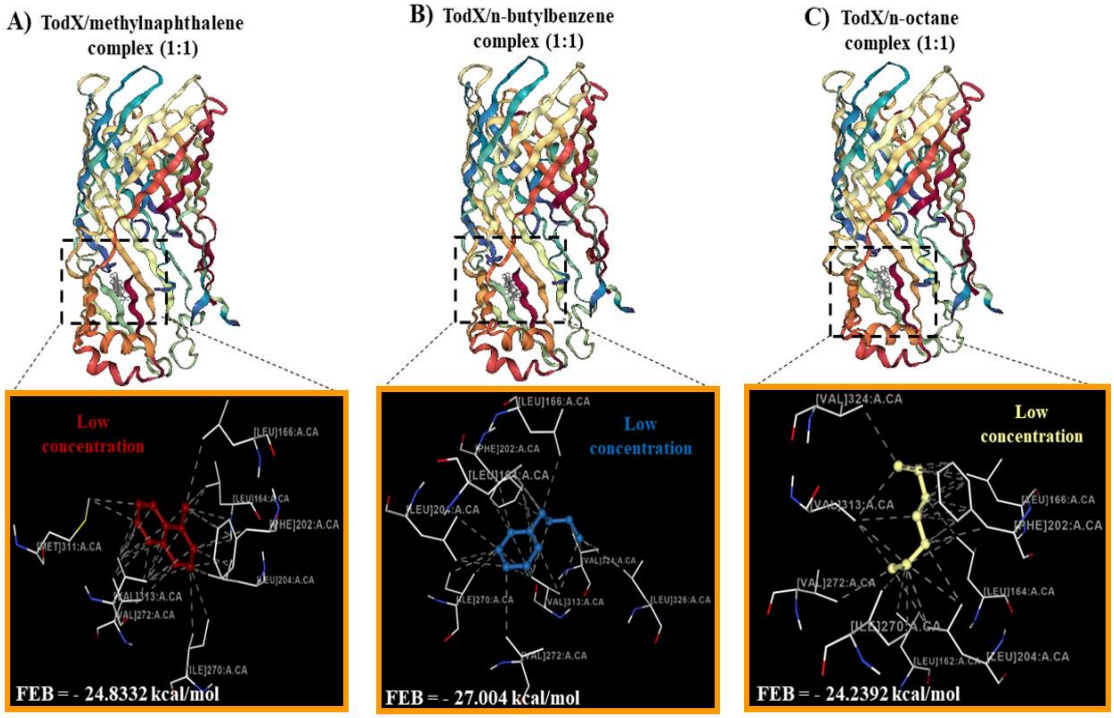


Figure 3. On the top, representation for each hydrocarbon/TodX docking complexes obtained for the best crystallographic-docking pose (RMSD < 2Å) using flexible molecular docking mimicking low hydrocarbon concentration (i.e., ligand-protein 1:1 proportion) in the TodX catalytic binding site 1. **A)** FEB (TodX/methylnaphtalene complex) = - 24.8332 kcal/mol with RMSD = 0.731Å, **B)** FEB (TodX/n-butylbenzene complex) = -27.004 kcal/mol with RMSD = 0.46Å, **C)** FEB (TodX/n-octane complex) = -24.2392 kcal/mol with RMSD = 0.72Å. On the bottom, representation of the corresponding 3D-lig-plot diagrams of interactions from the docking complex showing multiple hydrophobic interactions (gray dotted lines). For this instance, a color-label is assigned for each hydrocarbon like methylnaphthalene (red), n-butylbenzene (blue) and n-octane (yellow). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

Following this idea, we carried out the same docking simulation experiments using the two graphene nanostructures like pristine-Gr and oxidized-Gr in order to theoretically verify the interaction behavior of these systems mimicking low-concentrations.[29-34] See **Figure 4**.

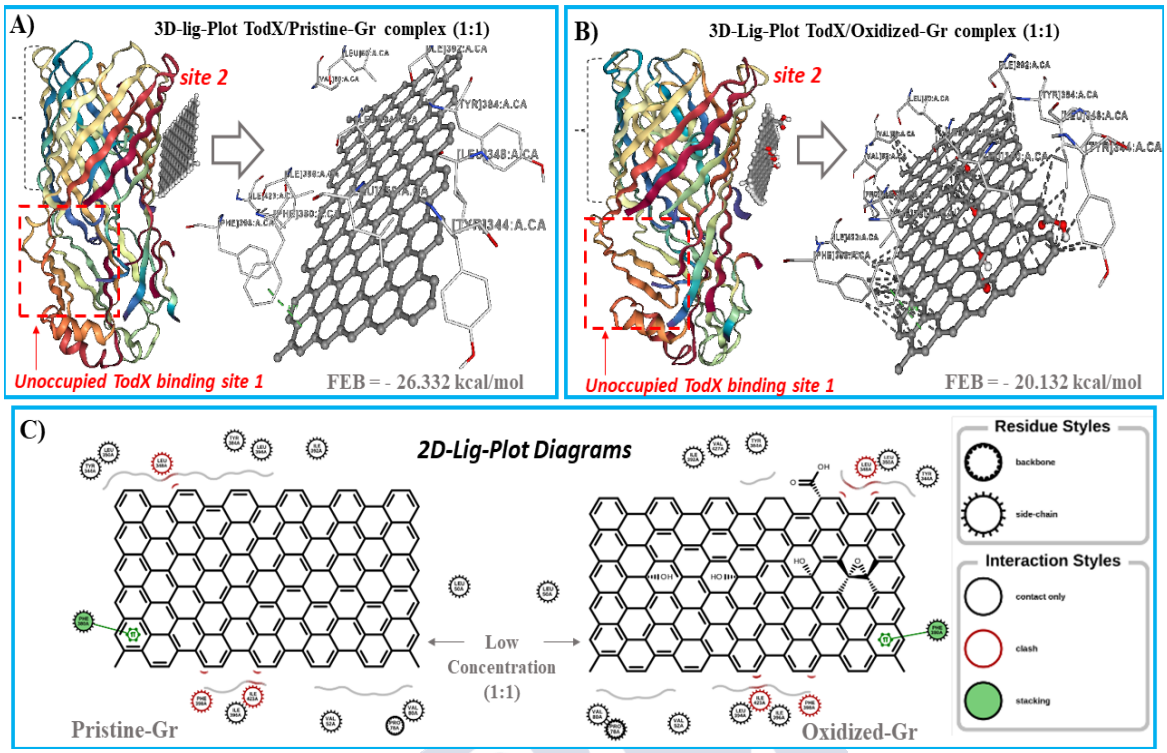


Figure 4. On the top, representation for each graphene nanostructure/TodX docking complexes obtained for the best crystallographic-docking pose (RMSD < 2Å) applying flexible molecular docking modeling the behavior of low graphene nanostructure concentrations (i.e., ligand-protein 1:1 proportion) in the TodX binding site 2, Please, note that the catalytic site 1 remains unoccupied for both carbon nanomaterials. **A)** FEB (TodX/pristine-Gr complex) = - 26.332 kcal/mol with RMSD = 1.431Å, **B)** FEB (TodX/oxidized-Gr complex) = -20.132 kcal/mol with RMSD = 1.406Å, in both panels are depicted on the right side, the corresponding 3D-lig-plot diagrams of the aforementioned docking complex showing multiple non-covalent hydrophobic interactions (gray dotted lines) highlighting the surrounding TodX-target amino acid residues from the site 2. On the bottom, **C)** Additional representation like 2D-lig-plot diagrams of interactions obtained from the cited docking complexes for pristine-Gr (left-side) and oxidized-Gr (right-site). Note the presence of multiple non-covalent hydrophobic interactions based on hydrophobic (N···C···C)-backbone together with side-chain contacts, a unique π–π stacking interaction with Phe 398-TodX (green) and several steric clashes (red). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

According to these results, we suggest that the interactions of graphene nanostructure in low-concentrations occurs in a different biophysical environment, involving lateral non-covalent hydrophobic contacts which, mainly affect the target residues from the extracellular domain binding site 2 according to (N⋯C⋯C)-backbone and side chains (A) interactions.[47-49] Then, considering that the biophysical environment of the interaction for the oil-derived hydrocarbons is placed in the TodX site 1 faraway of the TodX site 2, we strongly suggest that the biodegradation ability of the TodX protein could be partially affected under co-exposure with pristine-Gr and oxidized-Gr in low-concentration.

In this context, to evaluate the communication efficiency between the TodX binding site 1 and site 2 we carried out the ENM approach considering the collective motions in the TodX residue network (refer **Figure 1C**). The ENM analysis, allows the representation of the communication strength that each TodX residue has over every neighbor's residue (*i.e.*, residues from site 1 ↔ residues from site 2) based on the linear response theory.[24-26] For this instance, the collective fluctuation of the anisotropic network is defined by the $3N \times 3N$ Hessian matrix (H) following **equation 5**:

$$C_{i,j} = \begin{bmatrix} C_{1,1} & C_{1,2} & \ldots\ldots\ldots\ldots & C_{1,N} \\ C_{2,1} & C_{2,2} & \ldots\ldots\ldots\ldots & C_{2,N} \\ . & . & \ldots\ldots\ldots\ldots & . \\ C_{N,1} & C_{N,2} & \ldots\ldots\ldots\ldots & C_{N,N} \end{bmatrix} \textbf{(5)}$$

Where elements ($1 \leq i, j \leq N$) are the second derivatives of the ENM network potential (U) from **equation 6**:

$$C_{i,j} = \begin{bmatrix} \partial^2 U/\partial X_i \partial X_j & \partial^2 U/\partial X_i \partial Y_j & \partial^2 U/\partial X_i \partial Z_j \\ \partial^2 U/\partial Y_i X_j & \partial^2 U/\partial Y_i \partial Y_j & \partial^2 U/\partial Y_i \partial Z_j \\ \partial^2 U/\partial Z_i X_j & \partial^2 U/\partial Z_i \partial Y_j & \partial^2 U/\partial Z_i \partial Z_j \end{bmatrix} \textbf{(6)}$$

Herein, we use the biophysical parameter Markov commute-time [C (i, j) ↔ C (j, i)][24-26] to represent the ability of a given structurally communicated TodX binding residues (i, j) to receive and send signals in both directions ($j \rightarrow i$ and $i \rightarrow j$) from the site 1 to site 2 and vice versa. See **Figure 5**.

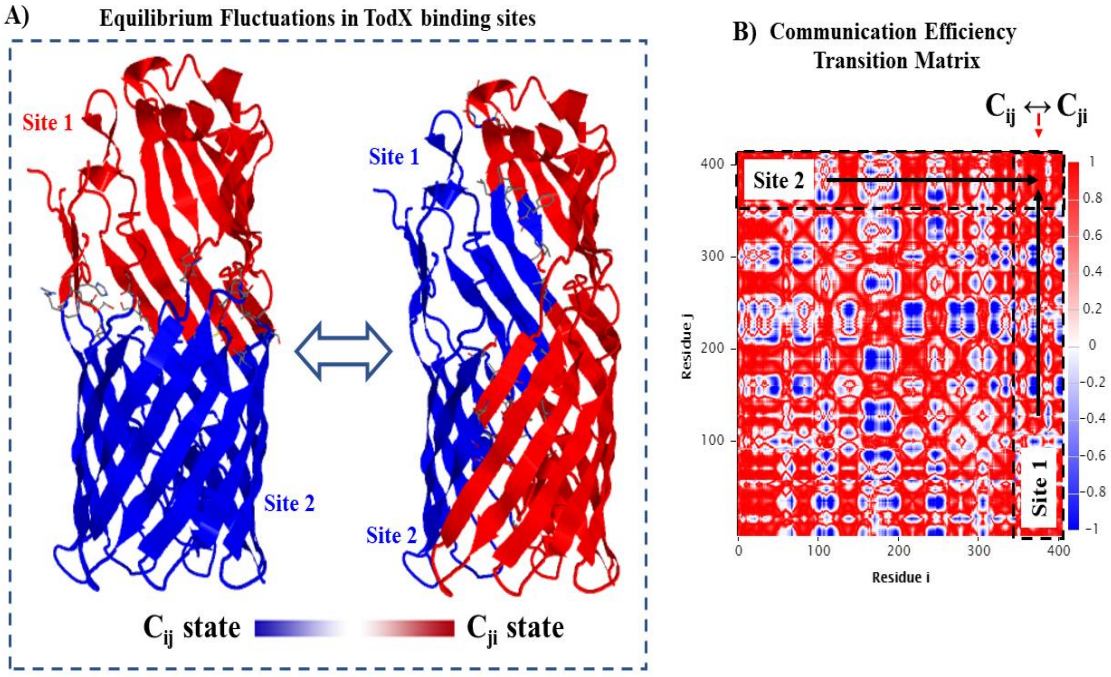


Figure 5. On the left, **A)** the 3D-cartoon representation of the inter-communication of TodX binding sites (site 1 and site 2) according to the different coupling states ([C (i, j) ↔ C (j, i)]) based on ENM approach. **B)** On the right, the corresponding two-dimensional communication efficiency transition matrix network-based on Markov commute time (C (i, j)), clustering receiving and communicating residues from the site 1 and site 2 of the TodX protein are represented by black-dotted rectangles. The color bar on the right side of the map indicates the pairs of TodX residues (i, j) that present fully correlated motions or strong correlations ($0.2 \leq (C_{ij} \leftrightarrow C_{ji}) \leq 1$) with labeled-color red indicating the same direction for the i and j residues-fluctuation, while the fluctuations based on anti-correlated motions of (i, j)-residue (i.e.: $C_{ij} \leftrightarrow C_{ji} \leq 0$) are colored in blue (opposite-direction for the fluctuation-motions of residues i and j) and the moderately correlated and uncorrelated ($C_{ij} \leftrightarrow C_{ji} \approx 0$) regions are labeled by light red and blue color, respectively. Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

The map of signal communication efficiency of TodX binding sites provides relevant information regarding the effective communication distances-based on the Markov commute time parameter (C (*i, j*)) for intra-segment C*α*-C*α* atomic distance fluctuations, which can be categorized like: *i)* fluctuations-based distances involving short segments (# TodX residues: 10–22), *ii)* fluctuations-based distance from medium segments (#

TodX residues: 23–26) and *iii)* fluctuations-based distance from long segments (TodX residues: 27–50).[50] Then, considering the results of the ENM approach we observe the presence of TodX-residues fluctuations involving coupled motion from medium to long distances based on the $C_{ij} \leftrightarrow C_{ji}$ parameter.[24-26] This approach suggests that both graphene nanostructures pristine-Gr that interacts with target residues (Phe398, Ile423, Ile396, Phe380, Val80, Leu50, Leu394, Leu350, Tyr384, Leu348, Tyr344) and the oxidized-Gr that interacts with target residues (Phe398, Ile423, Phe380, Ile396, Pro78, Val52, Val80, Leu50, Leu394, Ile392, Leu350, Tyr384, Leu348, Tyr344) in the extra-cellular site 2 (Refer to **Figure 1**). These fluctuations can induce potential perturbations in the biophysical environment of the site 1, where the main TodX catalytic residues are placed, and consequently could affect the biodegradation of the oil-derived hydrocarbon (n-butylbenzene, n-octane and methylnaphthalene) despite under low-concentration (1:1 protein-ligand proportion).[24-26]

The use of computational tools has been recommended and recognized by major regulatory agencies including the Organization for Economic Cooperation and Development (OECD, 2009)[51-54] and the International Organization for Standardization (ISO/TC 229, 2011)[51-54] based on the importance of developing alternative methods in nano-ecotoxicology.[51-54] The present docking approach try to offers a conceptual vision on the co-exposure phenomena, which are largely ignored from the theoretical point of view in nanoecotoxicological evaluations. The most studies address these environmental problems are carried out by performing predictive NanoQSAR models (Nano-Quantitative-Structure-Activity-Relationships).55 While mechanistic studies-based on multiple ligand-protein interactions explaining thermodynamic aspects of the carbon nanomaterial (in high-concentrations) remain unexplored until the present.

Following this idea, we carried out the simulation experiments considering simultaneous docking interactions, in order to theoretically model co-exposure conditions in high concentrations of the oil-derived hydrocarbons forming mixtures in the absence and the presence of detoxifying TodX protein and both graphene nanostructure (pristine-Gr and oxidized-Gr). Herein, the theoretical ecotoxicological response will depend on the ligand concentration (*i.e.,* ligand proportion stablished) and on the affinity-based energy difference for the ligands bound and/or unbound with the TodX protein. Following this idea, first, we will focus on the bound state in the presence of TodX protein and the graphene nanostructures mimicking co-exposure with oil-derived hydrocarbon in overload conditions.[33,34] As mentioned above the TodX protein allows the lateral

diffusion of hydrophobic substrates through the previously identified site 1 under normal conditions.[9] In this regard, the crystal structure of TodX protein from *Pseudomonas putida* shows that the opening in the wall of the beta-barrel is conserved composed by a long hydrophobic tunnel, which mediate the substrate passage from the extracellular environment, through a polar lipopolysaccharide layer and, through of the lateral opening in the barrel wall (embedded-membrane site 1), inside the lipid bilayer allowing the hydrophobic substrate diffusion into the periplasm.[9,10] However, under high oil-derived hydrocarbon concentrations and/or co-exposure conditions of interactions the biotransformation capacity is severely compromised, due to the saturation processes in the main catalytic residues composing the site 1, increasing the likelihood of inducing the loss of detoxifying ability linked to TodX channel nanotoxicity, the intensity of the response will depend on the hydrocarbon concentration and on the energy difference for a given hydrocarbon to be bound (*b*) or unbound (*unb*) with the TodX-protein.[27,33,34] Herein, considering the complexity of the problem of high hydrocarbon concentrations interacting with the TodX, and regarding that the total affinity (ΔG_{bind}) of the oil-derived hydrocarbons by TodX (catalytic site 1) can be modulated in the presence of carbon nanomaterials like graphene nanostructures, we carried out a molecular docking approach-based on Boltzmann multiplicity states.[27,33,34] For this instance, we can define the hydrocarbon overload interacting with TodX protein like a Boltzmann problem involving TodX channel nanotoxicity (*i.e.*, TodX site 1 saturation-induced by oil-derived hydrocarbons: methylnaphthalene, n-butylbenzene, n-octane and the graphene nanostructures: pristine-Gr and oxidized-Gr). Where the unbound state of oil-derived hydrocarbons and the graphene nanostructures in solution has a total unbound energy like $\varepsilon_{unb_{total}} = \sum\left(\sum\varepsilon_{unb\_hydrocarbons} + \sum\varepsilon_{unb\_graphenes}\right)$ and its corresponding energy $\varepsilon_{b_{total}} = \sum\left(\sum\varepsilon_{b_{hydrocarbons}} + \sum\varepsilon_{b_{graphenes}}\right)$, when forming the docking complexes with the *TodX*-receptor. Then, we can represent the Boltzmann problem-based distribution like an enumeration of multiple substrates (ligands) of docking interactions. See **Figure 6**.

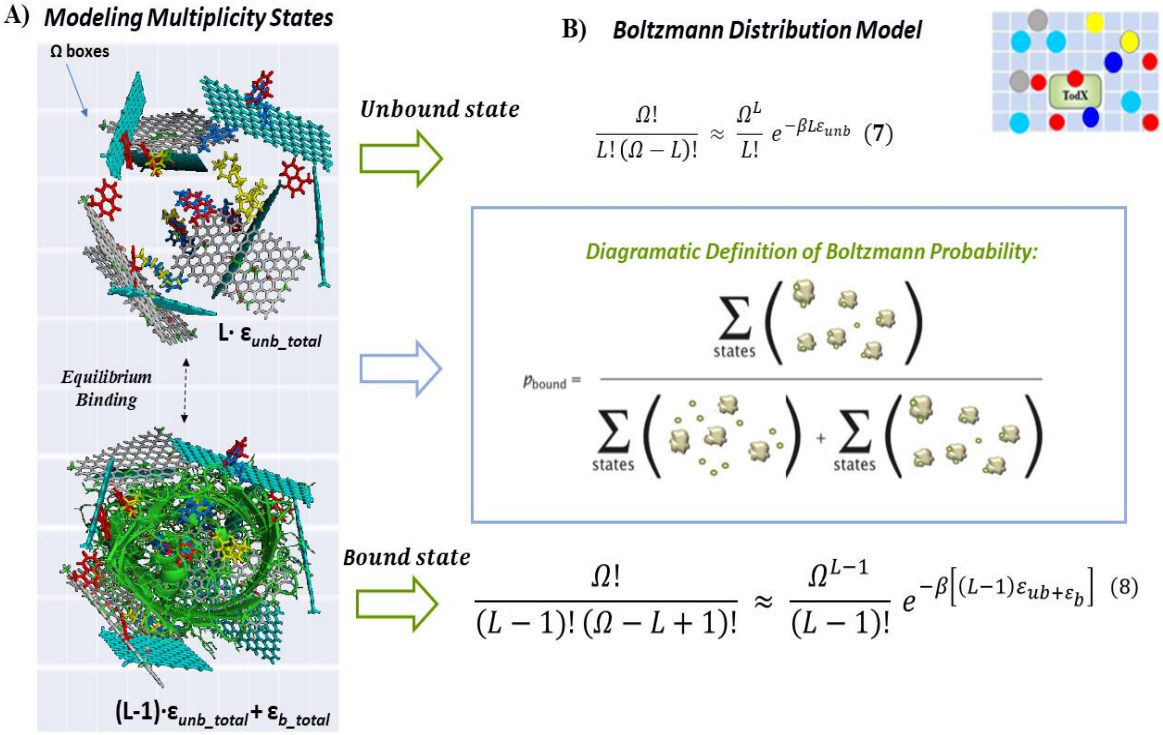


Figure 6. On the left, **A)** Representation of molecular docking approach based- Gibbs distribution under co-exposure equilibrium binding in multiplicity states,[33,34] showing the heterogeneous multicomponent docking complex formed by the oil-derived hydrocarbons like methylnaphthalene (red), n-butylbenzene (dark blue), n-octane (yellow), TodX protein (green), and the graphene nanostructures pristine-Gr (light blue) and oxidized-Gr (gray) in the unbound and bound state. On the right, **B)** Diagrammatic definition of the Gibbs distribution or Boltzmann probability (ρ_{bound}) of the system through several possible states. Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

Then, we can write the Boltzmann probability-based multiplicity states to represent ligand overload conditions like oil-derived hydrocarbons interacting with TodX and graphene nanostructures following the **equation 9**:

$$\rho_{bound} = \frac{e^{\frac{-\varepsilon_b}{k_B T}} \frac{\Omega^{L-1}}{(L-1)!} e^{-(L-1)\frac{\varepsilon_{unb}}{k_B T}}}{\frac{\Omega^L}{L!} e^{\frac{-L\cdot\varepsilon_{unb}}{k_B T}} + e^{\frac{\varepsilon_{unb}}{k_B T}} \frac{\Omega^{L-1}}{(L-1)!} e^{-(L-1)\frac{\varepsilon_{unb}}{k_B T}}} = \frac{\left(\frac{L}{\Omega}\right) e^{-\Delta\varepsilon/k_B T}}{1 + \left(\left(\frac{L}{\Omega}\right) e^{-\Delta\varepsilon/k_B T}\right)} \quad (9)$$

Herein, *L* is a function of the ligand concentration (c) under co-exposure conditions with fixed volume (V = 1 nm^3) in the boxes (Ω). In the present study the obtained multicomponent docking complex exhibits macrocanonical collectivity ensemble due to in the modeling systems the chemical potentials, volume (*V*) and temperature (*T*) were fixed. Besides, under thermodynamic equilibrium binding we can model these systems in terms of concentration like $c = \frac{L}{V}$ in co-exposure condition with high ligand concentrations. To this end, we can write the ρ_{bound} in term of ligand concentration (*c*) and equilibrium constant (K_{eq}) following the **equation 10** and **11**:

$$\rho_{bound} = \frac{\left(\frac{c}{c_0}\right) e^{\frac{-\Delta\varepsilon}{k_B T}}}{1 + \left(\frac{c}{c_0}\right)} \qquad (\mathbf{10})$$

$$\rho_{bound} = \frac{\left(\frac{c}{K_{eq}}\right)}{1 + \left(\frac{c}{K_{eq}}\right)} \qquad (\mathbf{11})$$

Where the $c_0 = \frac{\Omega}{V}$, the $\Delta\varepsilon = \varepsilon_b - \varepsilon_{unb} < 0$; the Boltzmann constant is represented by k_B, the K_{eq} has concentration units and can be expressed as $K_{eq} = c_0 \cdot e^{\frac{\Delta\varepsilon}{kT}}$, where if the $\Omega = 1cm^3$ and c_0 = 0.6 M. The Boltzmann probability distribution (ρ_{bound}) shows that states with lower energy will always be more likely to be occupied than states with higher energy. In addition, the ρ_{bound} provides relevant information about the quantitative relationship between the probabilities of the two bounded states (*i.e.,* concomitant interactions of the cited ligands at the same biophysical environment in TodX protein).[33,34] Despite these cases (two bounded-states) have similarities, it is important to distinguish them, since they generalize differently when crucial assumptions are changed. For this instance, under thermodynamic equilibrium binding regarding both energy exchange and ligand exchange, the thermodynamic requirement, fit with a relaxed fixed-composition obtaining a grand-canonical ensemble.[33,34]

According to our results we show that the co-exposure of oil-derived hydrocarbon in overload condition with *Pseudomonas putida* protein (*TodX*) and graphene nanostructures follows a complex multisite binding case. The aforementioned model of multicomponent binding equilibria has been prviously studied, by several authors.[33,34] These theoretical results are included in a theory of Heterogeneous Monod-Wyman-

Changeux (HMWC) multicomponent binding systems.[54,56-58] In this context, we can study the heterogeneous multicomponent binding equilibrium constant by modeling the fractional occupancy or TodX-saturation parameter (δ) as a measure of the affinity constant (*Ki*) to represent the binding of different ligands (oil-derived hydrocarbon and graphene nanostructures) interacting with multiple TodX binding sites (site 1, site 2,..., $site_{(n)}$), also considering non-predicted binding sites. In this regard, the co-exposure scenario and high concentration conditions for multiple ligands (m) (*i.e.*, all the oil-derived hydrocarbon and graphene nanostructures),[54-58] simultaneously interacting with non-identical TodX binding sites (n_k) can be represented according to the **equation 12**:

$$\delta_{TodX_{site}} = \frac{1}{n}\frac{\Pi_{k=1}^{m}\left(\sum_{i=1}^{n_k}\left(|X_k|K_i^R\Pi_{j\neq i}\left(1+c_j|X_k|K_j^R\right)\right)\right)+L\Pi_{k=1}^{m}\left(c_i|X_k|K_i^R\Pi_{j\neq i}\left(1+c_j|X_k|K_j^R\right)\right)}{\Pi_{k=1}^{m}\Pi_{i=1}^{n_k}(1+|X_k|K_i^R)+L\Pi_{k=1}^{m}\Pi_{i=1}^{n_k}(1+c_i|X_k|K_i^R)} \quad (12)$$

Where the different ligand-affinities are represented by the equilibrium constants (K_i, K_j). The parameter L represent an isomerization constant of the TodX protein take on pseudo-allosteric behavior from $L = \frac{|T_0|}{|R_0|}$, which describe the equilibrium between the high-affinity relaxed state (R_0) and the low-affinity tensor state (T_0) from TodX-binding sites. In this context, the different affinity is represented in the high-affinity R-state (K_i^R) for all the evaluated ligands and the c parameter is associated to how much the equilibrium between T and R states changes under co-exposure conditions.[54,56-59]

Next, we evaluate the energetic contributions for the total affinity (ΔG_{bind}) considering that the ligands binding interactions occur in the high-affinity relaxed conformation state (R_0) of TodX. For this instance, the cases in the absence of the TodX protein will be separately discussed by DFT-simulation. Then, the theoretical results on the formed heterogeneous multicomponent docking complexes are presented below following the order: ***system I***: methylnaphthalene + pristine-Gr + oxidized-Gr, ***system II***: n-butylbenzene + pristine-Gr + oxidized-Gr, ***system III***: n-octane + pristine-Gr + oxidized-Gr. See the first system in **Figure 7.**

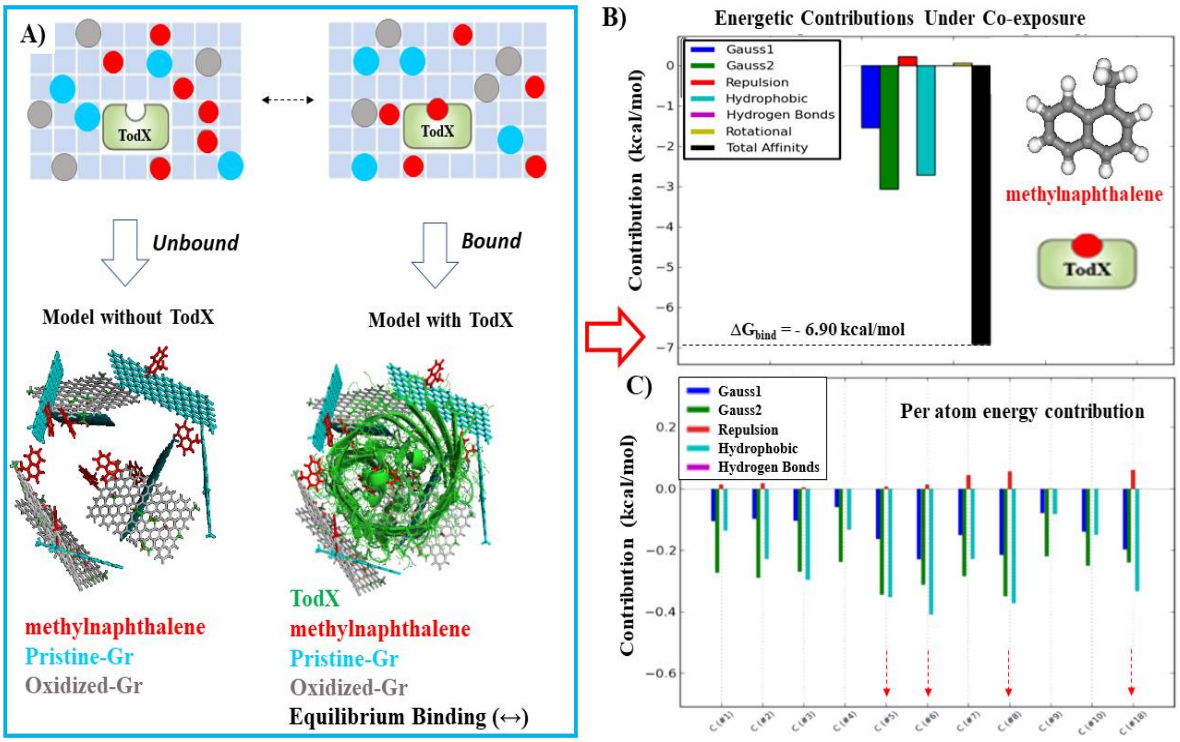


Figure 7. On the left side, **A)** Representation of the equilibrium binding-based on the heterogeneous multicomponent docking complex for the **system I**: methylnaphthalene (red) + pristine-Gr (light blue) + oxidized-Gr (gray) in high concentration of these ligands. Two co-exposure conditions are represented-like: i) without detoxifying TodX protein and ii) with detoxifying TodX protein. On the upper right side, **B)** Graphical breakdown of the methylnaphthalene binding energy contributions to the total affinity (ΔG_{bind}) with the corresponding values (kcal/mol) under co-exposure conditions with graphene nanostructures pristine-Gr and oxidized-Gr. On the lower right side, **C)** Graphical representation based on per atom-energy contribution for each methylnaphthalene-atom under co-exposure conditions with graphene nanostructures. For this instance, the methylnaphthalene-atoms with the best contribution to the total affinity (ΔG_{bind}) are highlighted (dotted-arrows labeled red) that fit with the light-blue bar for hydrophobic C-atom (C#5, C#6, C#8, and C#18). Herein, the symbol (#) x-axis it was used just for atom-labeling position purposes in the methylnaphthalene structure. Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

In the case of the methylnaphthalene, the results show that the most relevant energetic contributions to the total affinity (ΔG_{bind}) are provided by a strongest Gaussian molecular mechanism distance-dependent forces (Gauss2) which represent London non-electrostatic attractive energy contributions, followed by the contribution of non-covalent hydrophobic interactions ΔG_{VdW}.[28,37,59] In this regard, the analysis of the maximum atom-energy contribution for the methylnaphthalene shows that the most relevant interacting atoms were the C-atom-labeling positions (C#5, C#6, C#8, and C#18) which provided the strongest hydrophobic interactions. Furthermore, we theoretically suggest that in co-exposure conditions, the presence of graphene nanostructures (pristine-Gr and oxidized-Gr) significantly affected the methylnaphthalene affinity by the detoxifying TodX-protein, according to the results obtained from the ***system I*** (methylnaphthalene + pristine-Gr + oxidized-Gr + TodX), where the total affinity ΔG_{bind} (methylnaphthalene) = - 6.90 kcal/mol compared with the high-affinity value obtained in low-concentration like 1:1 protein-ligand proportion with ΔG_{bind} (methylnaphthalene/Todx) = -24.8332 kcal/mol (please, refer to **Figure 3A**).

Afterward, we present the ***system II:*** n-butylbenzene + pristine-Gr + oxidized-Gr in the **Figure 8.**

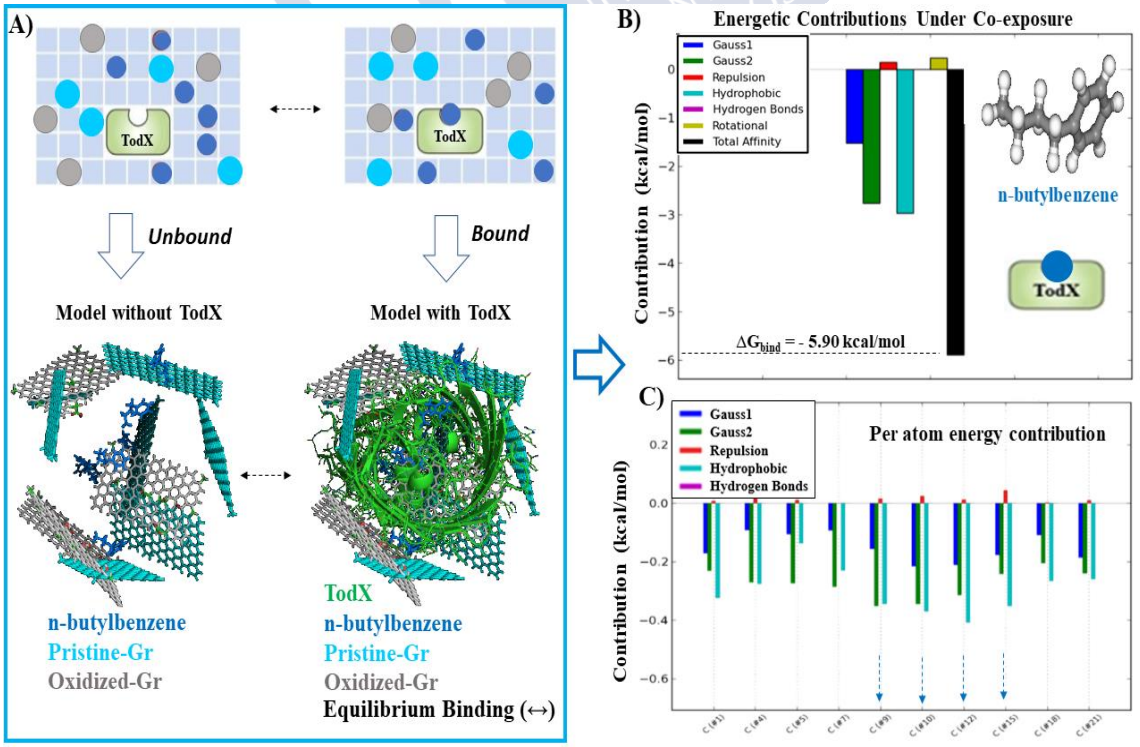


Figure 8. On the left side, **A)** Representation of the equilibrium binding-based on the heterogeneous multicomponent docking complex for the **system II**: n-butylbenzene (dark

blue) + pristine-Gr (light blue) + oxidized-Gr (gray) in a high concentration of these ligands. Two co-exposure conditions are represented-like: i) without detoxifying TodX protein and ii) with detoxifying TodX protein. On the upper right side, **B)** Graphical breakdown of the n-butylbenzene binding energy contributions to the total affinity (ΔG_{bind}) with the corresponding values (kcal/mol) under co-exposure conditions with graphene nanostructures pristine-Gr and oxidized-Gr. On the lower right side, **C)** Graphical representation based on per atom-energy contribution for each n-butylbenzene-atom under co-exposure conditions with graphene nanostructures. For this instance, the n-butylbenzene-atoms with the best contribution to the total affinity (ΔG_{bind}) are highlighted (dotted arrows labeled dark-blue) that fit with the light-blue bar for hydrophobic C-atom (C#9, C#10, C#12, and C#15). Herein, the symbol (#) x-axis it was used just for atom-labeling position purposes in the n-butylbenzene structure. Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

The simulated co-exposure ***system II****:* n-butylbenzene + pristine-Gr + oxidized-Gr show that the most dominant energy contributions for the total affinity (ΔG_{bind}) = -5.90 kcal/mol were provided from non-covalent hydrophobic Van der Waals interactions ΔG_{vdW} > Gauss2 attractive energy contributions. In this regard, the most relevant interacting atoms were the C-atom-labeling positions (C#9, C#10, C#12, and C#15) which provided the strongest hydrophobic interactions and London non-electrostatic attractive energy contributions. [54,56-59] Likewise, we note that the n-butylbenzene total affinity by the metabolizing TodX-protein significantly decreases in co-exposure conditions and high concentration of graphene nanostructures (pristine-Gr and oxidized-Gr) compared with the affinity reference value-based low-concentration of the n-butylbenzene/TodX complex with ΔG_{bind} = -27.004 kcal/mol (refer to **Figure 3B**). Next, following the same criteria, the third ***system III*** like n-octane + pristine-Gr + oxidized-Gr was evaluated. See **Figure 9**.

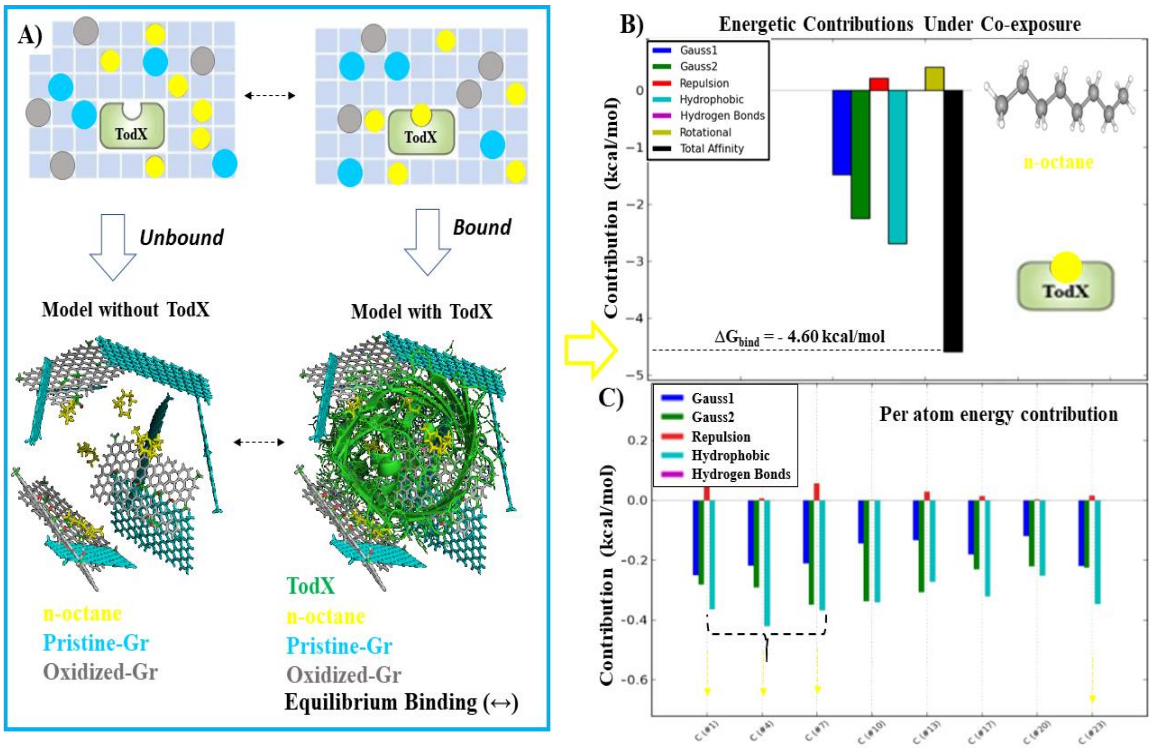


Figure 9. On the left side, **A)** Representation of the equilibrium binding-based on the heterogeneous multicomponent docking complex for the **system III**: n-octane (yellow) + pristine-Gr (light blue) + oxidized-Gr (gray) in a high concentration of these ligands. Two co-exposure conditions are represented-like: i) without detoxifying TodX protein and ii) with detoxifying TodX protein. On the upper right side, **B)** Graphical breakdown of the n-octane binding energy contributions to the total affinity (ΔG_{bind}) with the corresponding values (kcal/mol) under co-exposure conditions with graphene nanostructures pristine-Gr and oxidized-Gr. On the lower right side, **C)** Graphical representation based on per atom-energy contribution for each n-octane-atom under co-exposure conditions with graphene nanostructures. For this instance, the n-octane-atoms with the best contribution to the total affinity (ΔG_{bind}) are highlighted (dotted arrows labeled yellow) that fit with the light-blue bar for hydrophobic C-atom (C#1, C#4, C#7 and C#23). Herein, the symbol (#) x-axis it was used just for atom-labeling position purposes in the n-octane structure. Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

In this concern, the total affinity (ΔG_{bind}) = - 4.5 kcal/mol obtained from the n-octane/TodX complex in co-exposure conditions with graphene nanostructures, kept the same pattern on relevant energy contributions like (ΔG_{VdW} > Gauss2). For this instance, the atoms with the maximum influence in the hydrophobic interactions were the C-atom-

labeling positions (C#1, C#4, C#7 and C#23) associated to a significant loss of the relative affinity compared with the simulated condition-based low-concentration with ΔG_{bind} = -24.2392 kcal/mol (refer to **Figure 3C**).

In general terms, as expected the non-covalent hydrophobic interactions leading the main contributions to the total binding affinity with a spontaneous thermodynamic process ($\Delta G_{bind} < 0$) of the three evaluated oil-derived hydrocarbons multicomponent docking systems (*I*, *II*, and *III*) in co-exposure conditions with graphene nanostructures (pristine-Gr and oxidized-Gr) by their chemical properties. [54,56-59] This fact, can be corroborate by the total absence of H-bond interactions with the TodX-protein for the three oil-derived hydrocarbons and, considering the influence of the graphene nanostructures evaluated, which present a well-known network of sp^2-C-atoms with high surface reactivity, which significantly contribute to weaken the affinity of the oil-derived hydrocarbons forming multicomponent docking systems (*I*, *II*, and *III*). In addition, we show that most of the interactions are based on hydrophobic (N···C···C)-backbone and side chain interactions with the catalytic residues from the TodX (site 1).

Then, to briefly show the mutual influence of each ligands composing the heterogeneous multicomponent docking complex in the binding equilibria (unbound and bond) all the energetic contributions under co-exposure conditions are depicted. See **Figure 10**.

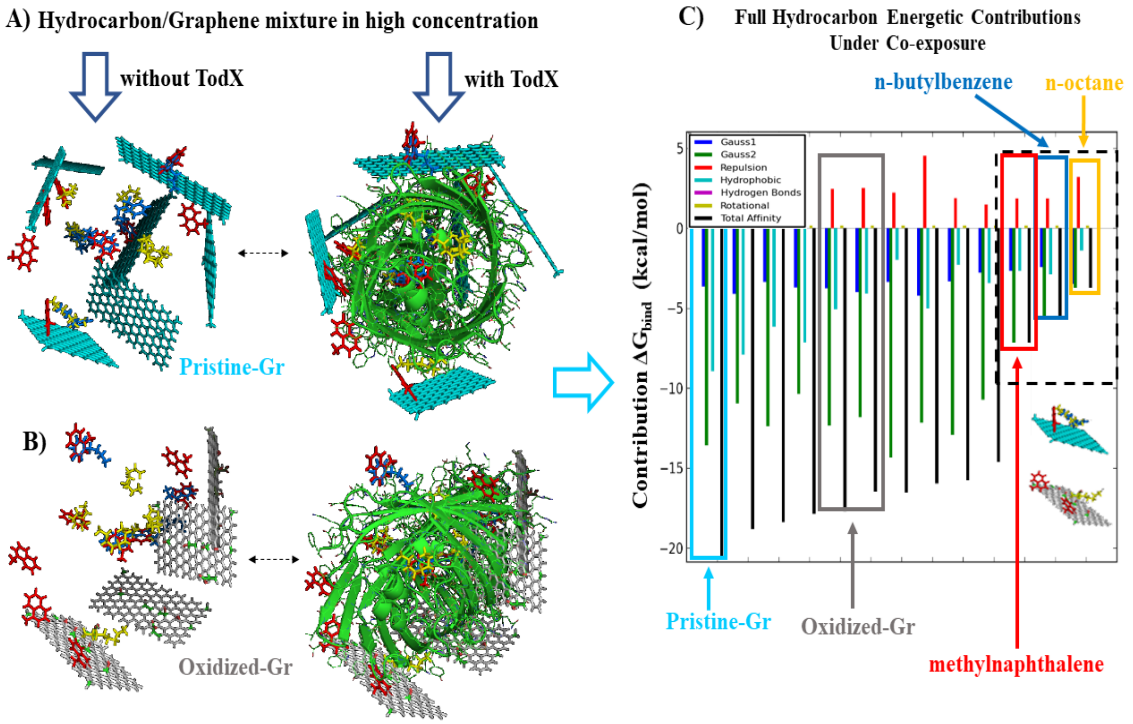


Figure 10. On the left side, **A)** Representation equilibrium binding-based Boltzmann model of the heterogeneous multicomponent docking complexes composed by the mixture of all the hydrocarbons interacting with pristine-Gr in the absence and presence of TodX. **B)** Heterogeneous multicomponent docking complexes composed by the mixture of all the hydrocarbons interacting with oxidized-Gr in the absence and presence of TodX. **C)** Graphical representation of all the energetic contributions to the total affinity (ΔG_{bind}) for each evaluated ligand in co-exposure conditions. Herein, the corresponding values of energetic contributions obtained for the mixture of all the oil-derived hydrocarbons are represented inside the black-dotted line rectangle like: methylnaphthalene (red), n-butyl benzene (blue), n-octane (yellow) and the graphene nanostructures like pristine-Gr (light blue) and oxidized-Gr (gray). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

For this instance, is easy to note that, in co-exposure conditions with multiple ligands the total affinity-based-binding free energy significantly decrease in all the docking systems following the order: pristine-Gr/TodX > oxidized-Gr/TodX > methylnaphthalene/TodX > n-butylbenzene/TodX > n-octane/TodX maybe because the graphene nanostructures could interact with oil-derived hydrocarbon resulting in a greater amount of TodX protein molecules are in an unbound state (*i.e.*, with unoccupied catalytic sites). [33,34, 54,56-59] On the other hand, the energetic contributions for the total affinity based on repulsive and rotational forces showed a similar behavior for the three oil-derived hydrocarbons/TodX complexes evaluated from the quantitative point of view. Besides, were considered as irrelevant for the stabilization of the multicomponent docking systems (*I*, *II*, and *III*) obtained under co-exposure conditions. [33,34, 54,56-59] These theoretical evidences fit with previous experimental results, which refer that carbon nanomaterials-like graphene's can mitigate the negative environmental impact of many organic compounds interact under co-exposure[15] in a given environmental compartment (water, soil). In this sense, we strongly support hypothesis that the presence of graphene nanostructure significantly reduces the affinity of the oil-derived hydrocarbons limiting its complete detoxification in the TodX-protein, and in this way, increasing its toxicological potential in environmental compartments unless remain adsorbed in the reactive surface of the graphene nanostructures.[15]

To verify the presence of physical-chemical adsorption processes from the graphene nanostructures (pristine-Gr and oxidized-Gr) interacting with the oil-derived hydrocarbons, a DFT study was carried out.[41-44] For this purpose, the electronic and structural properties of the interacting systems, previously defined on the experiment groups (**Group #1**: and **Group #2**), were analyzed and the main results are presented in the **Table 1**.

Table 1. DFT results of graphene nanostructures interacting with oil-derived hydrocarbons, where E_b, d, and ΔQ represent the binding energy, minimum distance, and charge transfer, respectively. Herein, the most stable configuration was obtained considering the stability of the formed complexes and denoted by the asterisk (*). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

DFT-Group	Ranking	DFT-systems	E_b (eV)	d(Å)	ΔQ (e^-)
#1	1	pristine-Gr + methylnaphthalene *	- 0.74	2.76	+ 0.05
#1	2	pristine-Gr + n-butylbenzene	- 0.39	3.44	+ 0.22
#1	3	pristine-Gr + n-octane	- 0.66	2.61	+ 0.51
#2	1	oxidized-Gr + methylnaphthalene *	-0.59	2.70	+ 0.06
#2	2	oxidized-Gr + n-octane	-0.44	2.18	+0.08
#2	3	oxidized-Gr + n-butylbenzene	-0.30	2.33	0.00

According to the obtained results, the binding energies (E_b) vary from -0.74 to -0.30 eV characterizing a spontaneous thermodynamic process in all the cases for both groups (Group #1 and Group #2). This fact has a paramount importance, because suggest high probability (*i.e.*, occurrence-based frequency) of interactions and binding affinity for the formed complexes. In the case of the oil-derived hydrocarbons interacting with pristine-Gr (Group #1), the best-ranked binding configuration (*) was found by the pristine-Gr + methylnaphthalene, which presents the highest modular binding-energy value and thermodynamic stability of $|E_b| = 0.74$ eV with an minimum distance of 2.76 Å. See the **Figure 11**.

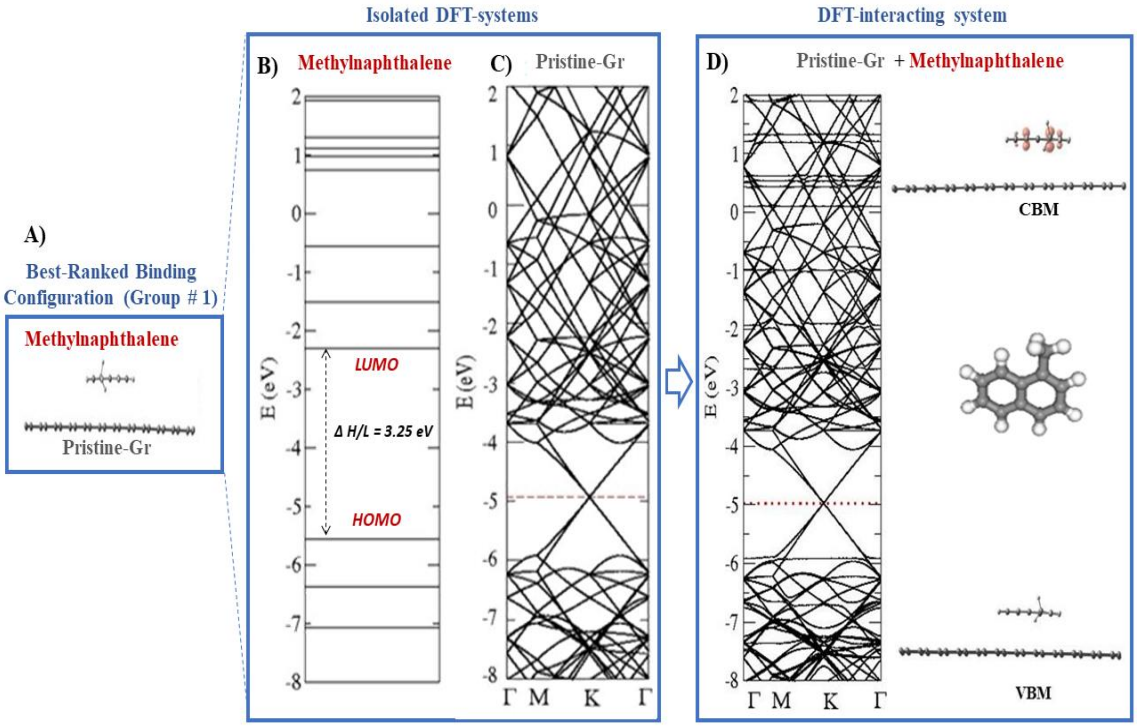


Figure 11. On the rightmost rectangle, **A)** Representation of the optimized structures of pristine-Gr (labeled-color gray) and methylnaphthalene (labeled-color red) from the best-ranked binding configuration in the Group #1. In the middle rectangle, are depicted the HOMO/LUMO energy levels and the electronic band structure for the isolated ligands: **B)** methylnaphthalene molecule (labeled-color red) and **C)** pristine-Gr (labeled-color gray), respectively. On the leftmost rectangle, **D)** represents the electronic band structure obtained from the DFT-interacting system formed by pristine-Gr + methylnaphthalene with its corresponding electronic charge density plot. An isosurface value of 0.00096 e^- /$Bohr^3$ was used to represents the charge density associated to conduction band minimum (CBM) and the valence band maximum (VBM). The atom-color labels were represented like: hydrogen-atoms (light-gray) and the carbon-atoms (dark-gray). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

As can be seen from **Figure 11A**, the methylnaphthalene molecule adopts a parallel position with respect to the pristine-Gr with a value of charge transfer of +0.05e-, showing that the electron transference occurs from the pristine-Gr to methylnaphthalene. According to this evidence, we can suggest the presence of a physical adsorption process, without significant modifications on the electronic band structure, around the Fermi region, after the pristine-Gr + methylnaphthalene interaction. Following this idea, we can

observe that, in the DFT-interacting system formed by pristine-Gr + methylnaphthalene exists a superposition (**Figure 11 D**) of the energy levels of the methylnaphthalene molecule (**Figure 11 B**) with the electronic band structure of the pristine-Gr (**Figure 11 C**), confirming that the interaction occurs through a physical adsorption regime. Also, the electronic charge density associated to CBM is only concentrated in the methylnaphthalene molecule (**Figure 11 D**). Further, for this particular isosurface value, no charge concentration was observed in the VBM (**Figure 11D**), which remain unperturbed for both ligands (methylnaphthalene and pristine-Gr) according to with the calculated binding energy (Eb = - 0.74 eV).

Regarding the main objective of this theoretical study the physical adsorption in the reactive surface of the pristine-Gr is a favorable characteristic toward potential applications as bioremediation strategies and environmental modeling of complexes mixtures.15 Because, it is well-known that the systems which adsorb through weak interaction patterns may present a reversible character during the complex formation, which allows the easy removal of the desired compounds (as oil-derived hydrocarbons) from the environmental compartments (water, soil), and later, the system could be efficiently reused in subsequent treatments, in the case of this nanostructure (pristine-Gr).15,60

By the other hand, the system formed by the oxidized-Gr + methylnaphthalene DFT-results were also identified like the best-ranked binding configuration (*) in the DFT-Group #2 with the highest modular binding-energy value and thermodynamic stability | Eb | = 0.59 eV compared with the other two DFT-complexes like oxidized-Gr + n-octane (| Eb | = 0.44 eV) and oxidized-Gr + n-butylbenzene (| Eb | = 0.30 eV). For this most stable configuration, the obtained value for the charge transfer, between an H-atom from the methylnaphthalene with an O-atom from oxidized-Gr, was + 0.06 e-. The electronic characteristics of the oxidized-Gr + methylnaphthalene system are shown in the **Figure 12**.

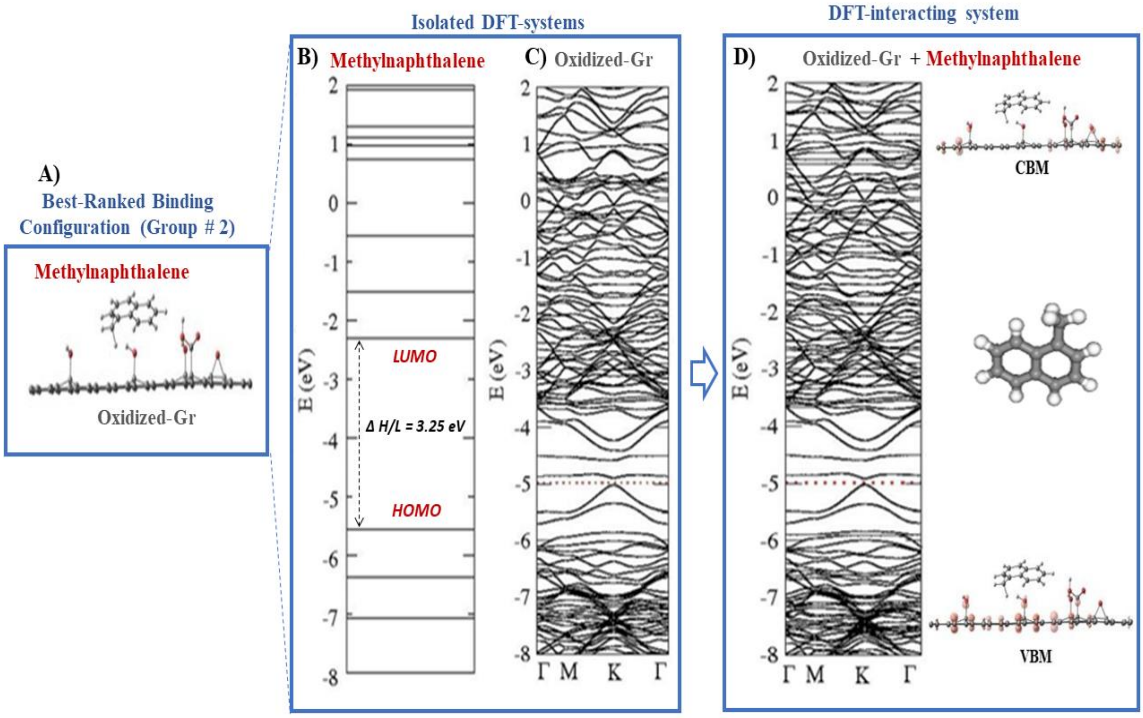


Figure 12. On the rightmost rectangle, **A)** Representation of the optimized structures of oxidized-Gr (labeled-color gray) and methylnaphthalene (labeled-color red) from the best-ranked binding configuration in the Group #2. In the middle rectangle, are depicted the HOMO/LUMO energy levels and the electronic band structure for the isolated ligands: **B)** methylnaphthalene (labeled-color red) molecule and **C)** oxidized-Gr (labeled-color gray), respectively. On the leftmost rectangle, **D)** represents the electronic band structure obtained from the DFT-interacting system formed by oxidized-Gr + methylnaphthalene with its corresponding electronic charge density plot. An isosurface value of 0.00096 $e^-/Bohr^3$ was used to represents the charge density associated to conduction band minimum (CBM) and the valence band maximum (VBM). The atom-color labels were represented like: oxygen-atoms (red), hydrogen-atoms (light-gray), and the carbon-atoms (dark-gray). Source adapted from Bentham Science Publisher, please visit: https://doi.org/10.2174/1568026620666200820145412 (accessed on 21 july 2020)

Herein, no significant differences were observed between the electronic properties of the isolated oxidized-Gr system (**Figure 12 C**), with respect to the oxidized-Gr + methylnaphthalene system (**Figure 12 D**). We note the presence of superposition of the energy levels of the methylnaphthalene molecule (**Figure 12 B**) with the electronic band structure of the oxidized-Gr after the interaction (**Figure 12 D**), suggesting that the interaction occurs also through a physical adsorption regime in the **DFT-Group #2,**

similar to the most stable system in the **DFT-Group #1**. As a difference, it is important to highlight that, in the case of the oxidized-Gr + methylnaphthalene system, the results of electronic charge density plot show that the charges were only concentrated on the oxidized-Gr nanostructure in both CBM and VBM edges (**Figure 12 D**).

In this context, the current results suggest that the pristine-Gr has a higher capacity to bind hydrophobic molecules like oil-derived hydrocarbon as compared with the oxidized-Gr, due to the π- π interactions between the graphene atoms and the aromatic rings of the adsorbed molecules. In addition, these theoretical evidences have been experimentally observed in previous works performed by Wang et al. (2014),[12] in which they show that the interaction affinity between hydrocarbons and carbon nanostructures are favored by the physical adsorption mechanism-based on hydrophobic π- π interactions between the sp^2-C atoms from aromatic rings of the molecules.[15,60]

An overview of our theoretical results strongly suggests that environmental mixtures of oil-derived hydrocarbon with carbon nanomaterials like graphene can potentially reduce the ecotoxicological impact of many other pollutants of high persistence in the environment. In this sense, the theoretically evaluated adsorption mechanism mimicking a well-recognized *Trojan-horse phenomenon* which leading most of the interaction processes for engineered organic nanomaterials (NM).[15,60] Nevertheless, still a long way to explore, since the physico-chemical mechanisms in the nanoscale, exhibit nonlinear behaviors (*i.e.*, nonlinear concentration-response relationships in the environmental compartments like water, soil and air), which depend on the global physical and chemical properties of a given mixture and the biological model used (molecules, proteins, cells, organism, ecosystems).

References

1. Barron, M. G. Ecological impacts of the Deepwater Horizon oil spill: implications for immunotoxicity. *Toxicol. Pathol*, **2012**, 40, 315-320.
2. Goldstein, B. D., Osofsky, H. J., & Lichtveld, M. Y. The Gulf oil spill. *New Eng. Journ. Med.* **2011**, 364, 1334-1348.
3. Beyer, J., Trannum, H. C., Bakke, T., Hodson, P. V., & Collier, T. K. (2016). Environmental effects of the Deepwater Horizon oil spill: a review. *Marine pollution bulletin*. **2016**, 110, 28-51.

4. Balba, M. T., Al-Awadhi, N., & Al-Daher, R. Bioremediation of oil-contaminated soil: microbiological methods for feasibility assessment and field evaluation. *Journ. Microbiol Meth*. **1998**, 32, 155-164.
5. Adams, G. O., Fufeyin, P. T., Okoro, S. E., & Ehinomen, I. Bioremediation, biostimulation and bioaugmention: a review. *Inter. Journ. Environm. Bioremed & Biod*, **2015**, 3, 28-39.
6. Raghavan, P. U. M., & Vivekanandan, M. Bioremediation of oil-spilled sites through seeding of naturally adapted Pseudomonas putida. *Intern biodeterioration & biodegradation*, **1999**, 44, 29-32.
7. Nwachukwu, S. C., James, P., & Gurney, T. R. Inorganic nutrient utilization by "adapted" Pseudomonas putida strain used in the bioremediation of agricultural soil polluted with crude petroleum. *Journ. Environm. Biology*, **2001**, 22, 153-162.
8. Belchik. A β-barrel outer membrane protein facilitates cellular uptake of polychlorophenols in *Cupriavidus necator*. *Biodegradation*. **2010**, 21, 431-439.
9. Van den Berg, B. Going forward laterally: transmembrane passage of hydrophobic molecules through protein channel walls. *Chembiochem*, **2010**, *11*, 1339-1343.
10. Mardani, G., Mahvi, A. H., Hashemzadeh-Chaleshtori, M., Naseri, S., Dehghani, M. H., & Ghasemi-Dehkordi, P. Application of Genetically Engineered Dioxygenase Producing *Pseudomonas putida* on Decomposition of Oil from Spiked Soil. *Jundishapur Journ. Nat. Pharm. Prod.* **2017**, *12* (3 (Supp)).
11. Kharisov, B. I., Dias, H. R., & Kharissova, O. V. Nanotechnology-based remediation of petroleum impurities from water. *Journ. Petrol. Sci. Eng.* **2014**, 122, 705-718.
12. Wang, J., Chen, Z., & Chen, B. Adsorption of polycyclic aromatic hydrocarbons by graphene and graphene oxide nanosheets. *Environm. Sci. Technol.* **2014**, 48, 4817-4825.
13. Iqbal, M. Z., & Abdala, A. A. Oil spill cleanup using graphene. *Environm. Sci. Poll. Res.* **2013**, 20, 3271-3279.
14. Robbins, W. K., & Hsu, C. S. Petroleum, Composition. *Kirk-Othmer Encyclopedia of Chemical Technology.* **2000**.
15. Naasz, S., Altenburger, R., Dana Kühne. Environmental mixtures of nanomaterials and chemicals: The *Trojan-horse* phenomenon and its relevance for ecotoxicity. *Journ. Sci. Tot. Env.* 2018, 635, 1170-1181.
16. Xu, X., Liu, Wenming., Tian, S., Wang, W., Qi, Qige., Jiang, Pan., Gao, Xinmei., Li, Fengjiao., Li, Haiyan., Yu, Hongwen. Petroleum Hydrocarbon-Degrading Bacteria for the Remediation of Oil Pollution Under Aerobic Conditions: A Perspective Analysis. *Front. Microbiol*. **2018**, 9, 2885.
17. Adebajo, M.O., Frost, R.L., Kloprogge, J.T., Carmody, O., Kokot, S. Porous materials for oil spill cleanup: a review of synthesis and absorbing properties. *J. Porous Mater*. **2003**, 10 , 159-170.
18. Martín de Lucía, I., Campos-Mañas, M.C., Agüera, A., Rodea-Palomares, I., G., Francisco Leganés, Fernández-Piñas, F., and Rosa, R. Reverse Trojan-horse effect

decreased wastewater toxicity in the presence of inorganic nanoparticles. *Environ. Sci.: Nano*. **2017**,4, 1273-1282.
19. Lun Hsiao, I., Kong Hsieh, Yi., Chu-Fang W., Chieh, Chen., Yuh-Jeen, Huang. Trojan-Horse Mechanism in the Cellular Uptake of Silver Nanoparticles Verified by Direct Intra- and Extracellular Silver Speciation Analysis. *Environ. Sci. Technol*. 2015, 49, 3813-3821.
20. Jiménez, J., Doerr, S., Martínez-Rosell, G., Rose, A. S., & De Fabritiis, G. DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics*, **2017**,33, 3036-3042.
21. Feinstein, W. P., & Brylinski, M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. *J Cheminform*. **2015**, *7*, 18.
22. Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc*, **2016**, 11, 905.
23. Vincent B. Chen, W. Bryan Arendall III, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson and David C. Richardson. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica*. **2010**, 66: 12-21.
24. Mitternacht, S., Berezovsky, I.N. Coherent Conformational Degrees of Freedom as a Structural Basis for Allosteric Communication. *Plos. Comput. Biol*. ***2011***, 7, e1002301.
25. Ozlem Keskin, Stewart R. Durell, Ivet Bahar, Robert L. Jernigan, and David G. Covell. Relating Molecular Flexibility to Function: A Case Study of Tubulin. *Biophys Journ*. **2002**, 83, 663–680.
26. Greener, J.G., Sternberg, M.J.E. AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. *BMC Bioinformatics*. **2015**, 16, 335.
27. Hearn, E.M., Patel, D.R., Bert van den Berg. Outer-membrane transport of aromatic hydrocarbons as a first step in biodegradation. *PNAS*, **2008**, 105, 8601-8606.
28. Trott, O., & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*, **2010**, 31, 455-461.
29. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem*. **1998**, *19*, 1639-1662.
30. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235-242.
31. Xie, Z. R., & Hwang, M. J. An interaction-motif-based scoring function for protein-ligand docking. *BMC bioinformatics*, **2010**, 11, 298.
32. Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B., Zhang, J., Bryant, S.H.

PubChem substance and compound databases. *Nucleic Acids Res*. **2016**, 44, 202–1213.
33. Feldman H.A. Mathematical Theory of Complex Ligand-Binding Systems at Equilibrium: Some Methods for Parameter Fitting. *Anal. Biochem*. **1972**, 48, 317-338.
34. Haiech, J, Gendrault, Y., Kilhoffer, M.C., Ranjeva, R., Madec, M., Lallement, C. A general framework improving teaching ligand binding to a macromolecule. *Bioch. Biophys. Act*. 2014, 1843, 2348–2355.
35. Seeliger, D., & de Groot, B. L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina., *J Comput Aided Mol Des. **2010**, 24*, 417-422.
36. Laskowski, R.A, Swindells, M.B. LigPlot+: Multiple LigandProtein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model*. **2011**, 51, 2778–2786.
37. Aoxiang Tao, Dong Xu, et al. ezCADD: A Rapid 2D/3D Visualization-Enabled Web Modeling Environment for Democratizing Computer-Aided Drug Design. : *J. Chem. Inf. Model*. **2019**, 59, 18−24.
38. da Silveira, C. H., Pires, D. E., Minardi, R. C., Ribeiro, C., Veloso, C. J., Lopes, J. C., Santoro, M. M. Protein cutoff scanning: A comparative analysis of cutoff dependent and cutoff free methods for prospecting contacts in proteins. *Proteins: Struct Funct Bioinf*. **2009**, 74, 727-743.
39. Rarey, M., Kramer, B., & Lengauer, T. Multiple automatic base selection: Protein–ligand docking based on incremental construction without manual intervention. *J Comput Aided Mol Des*, **1997**, *11*: 369-384.
40. Shoichet, B. K. Virtual screening of chemical libraries. *Nature.* **2004**, 432, 862-865.
41. Soler, J. M., Artacho, E., Gale, J. D., García, A., Junquera, J., Ordejón, P., & Sánchez-Portal, D. The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter*. **2002**, 14, 2745.
42. Kohn, W., & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev*.**1965**, 140, 1133.
43. Perdew, J. P., & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*. **1981**, 23, 5048.
44. Troullier, N., & Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, **1991**, 43, 1993.
45. Adipah, S. Introduction of Petroleum Hydrocarbons Contaminants and its Human Effects. *Journ. Env. Sci. Pub. Health.* **2019**, 3, 001-009.
46. Das, N., Chandran, P. Microbial Degradation of Petroleum Hydrocarbon Contaminants: An Overview. *Biotechnol. Res. Inter*. **2010**, 10, 1-13.
47. Gray, J. S. Biomagnification in marine systems: the perspective of an ecologist. *Mar Pollut Bull*. 2002, 45, 46-52.
48. Riisgaard, H.U., Hansen, S. Biomagnification of mercury in a marine grazing food-chain: algal cells *Phaeodactylum tricornutum*, mussels *Mytilus edulis* and flounders *Platichthys flesus*studied by means of a stepwise-reduction-CVAA method. *Mar. Ecol. Prog. Ser*. **1990**, 62, 259-270.

49. Strandberg, B., Bandh, C., Van Bavel B., Bergqvist P. A. , Broman, D., Naef, C., Pettersen, H., Rappe, C. Concentrations, biomagnification and spatial variation of organochlorine compounds in a pelagic food web in the northern part of the Baltic Sea. *Sci. Total Environ*. **1998**, 217, 143-154.
50. Ikeda, K.; Hirokawa, T.; Higo, J.; Tomii, K., Protein-segment universe exhibiting transitions at intermediate segment length in conformational subspaces. *BMC. Struct Biol* **2008,** *8*.
51. Patlewicz, G., Worth, A. P., and Ball, N. Validation of Computational Methods. *Adv Exp Med Biol*. 2016, *856*, 165-187.
52. OECD Principles for the validation, for regulatory purposes of (Quantitative) Structure Activity Relationship Model. http: www.oecd.org/(accessed 26/08/**2019**).
53. ISO/TC 229 Nanotech., **2011**.
54. Benko, H. ISO Technical Committee 229 Nanotechnologies. Metrology and Standardization of Nanotechnology: Protocols and Industrial Innovations. **2017,** 527-34039.
55. Ghosh, S., Ojha, P. K, Roy, K. Exploring QSPR modeling for adsorption of hazardous synthetic organic chemicals (SOCs) by SWCNTs. *Chemosphere*, **2019**, 228, 545-555.
56. Haiech, J., Vallet, B., Aquaron, R., Demaille, J.G. Ligand binding to macromolecules: determination of binding parameters by combined use of ligand buffers and flow dialysis; application to calcium-binding proteins, *Anal. Biochem.* **1980**,105, 18–23.
57. J.E. Fletcher, J.E., Spector, A.A., Ashbrook, J.D. Analysis of macromolecule—ligand binding by determination of stepwise equilibrium constants, *Biochem*. **1970**, 9, 4580–4587.
58. Klotz, I.M., Hunston, D.L. Mathematical models for ligand–receptor binding. Real sites, ghost sites, *J. Biol. Chem*. **1984**, 259, 10060-10062.
59. Klotz, I.M. Ligand–receptor interactions: facts and fantasies, *Q. Rev. Biophys.* **1985**, 18, 227—259.
60. Konkoli, Z. Safe uses of Hill's model: an exact comparison with the Adair—Klotz model, *Theor. Biol. Med. Model*. **2011**, 8, 10.
61. Sánchez-Linares, I., Pérez-Sánchez, H., Cecilia, J.M., García, J, M. High-throughput parallel blind virtual screening using BINDSURF. *BMC Bioinformatics*. **2012**, 13, 13.
62. Jauris, I.M., Matos, C. F. , Saucier C., Lima, E. C., Zarbin, A. J. G., Fagan, S. B., Machado, F. M. , Zanella, I. Adsorption of sodium diclofenac on graphene: a combined experimental and theoretical study. *Phys. Chem. Chem. Phys*. **2016,** 18, 1526-1536.

Chapter 11.

Concluding remarks

In the present thesis, several protein-ligand interaction systems with potential relevance to the fields of Medicinal Chemistry, Material Sciences, Nanomedicine, and Environmental Health Sciences were investigated. A combination of theoretical approaches, mainly based on thermodynamic structure-based molecular docking and experimentally validated techniques were used for conventional ligands belonging to the class 1 (as pharmaceutical drugs). While for the ligands belonging to the class 2 as carbon nanoparticles (or nanomaterials) mainly was emphasized on the importance of theoretical and computational methods.

Then, the main contributions of the present work focus on the importance and novel applications of state-of-the-art thermodynamic methods such as theoretical perturbation approaches to study conformational properties and allosteric signal propagation in proteins in the unbound and bound states, fractal approaches to solve protein-ligand complementarity problems, as well as, predictive models based on quantitative structure-binding-toxicity relationships at the nanoscale (nano-QSBR/QSTR models) toward efficient drug-discovery and systems with desired low-toxicity (including nanotoxicity). Then, by integrating the theoretical and empirical mechanistic findings, the conclusions from this work can be summarized as follows:

- The structural effects and binding properties of two betablockers (acebutolol and propranolol) on fibrinogen were evaluated The *in silico* results showed that the free energy of binding values for the obtained docking complexes (acebutolol/fibrinogen and propranolol/fibrinogen) were very close and following a spontaneous thermodynamic process (negative FEB values). Furthermore, the obtained theoretical results show that the mechanism of docking for the propranolol and acebutolol is mainly-based on non-covalent hydrophobic interactions with the fibrinogen E-region binding-site. Showing that the propranolol and acebutolol are able to induce significant changes in the inter-residue communication and flexibility-properties of residue network of fibrinogen E-region by inducing local-perturbations in allosteric residues. Likewise, the aforementioned ligands can affect the fluctuation patterns-based elastic normal

modes of flexibility by decreasing the biophysical parameters: frequency and collectivity degree in the funnel-shaped hydrophobic cavity of E-region. The obtained theoretical results fit very well with previous experimental evidences obtained by our group by using ITC and SAXS analysis. Besides, these results could have relevant importance in the study and understanding the coagulation processes under unbound and bound states where the binding of thrombin at E-domain and subsequently fibrinogen polymerization are crucial aspects.

- The influence of different BTS-ligand conformation binding modes under the interaction with protein fibrinogen was evaluated by using molecular docking simulations. The computational results showed that the Gibbs free energy of the obtained docking complexes (BTS *poses*_1-9 at the fibrinogen E-region) were very close in most of the cases, and indicating also that their formation is related to spontaneous thermodynamic processes. The differences detected in the FEB values for the nine BTS-poses may be explained based on the different composition and binding properties of the interacting residues present in the biophysical environment of the fibrinogen binding-site (E-region). Besides, the mechanism of docking for the different BTS-binding poses appears to be mainly due to non-covalent hydrophobic interactions, like Π-Π stacking ones, and associated to H-bonding, which together contributes to the stability of the formed complexes. Further experimental results from ITC confirmed the results from Gibbs free energy of binding and the interactions involved in the process. Also, suggesting that the different BTS-binding poses can induce simultaneous modulation (changes) of several target residues (*i.e.* of effector and allosteric residues), based on the results of the local perturbation-response scanning maps, by affecting the inter-residue communication and flexibility-properties of the fibrinogen E-region residue network, according to the most negative FEB values of interaction (BTS-binding poses 1 to 3) obtained. Likewise, we show that all the BTS-binding poses can directly affect the normal modes of flexibility by decreasing the collectivity degree in the fibrinogen binding-site (residues of E-region). Regarding the LPRS maps image-processing combining 2D-fast Fourier transform and fractal approaches, it revealed significant differences in the frequency domain signals of the BTS-interaction attributed to small conformational changes through the fibrinogen chain. Lastly, these theoretical

results were confirmed by using ITC and SAXS measurements, that not only validated the presence of fibrinogen-BTS conformational binding interactions, but also revealed a strong dependence on the BTS-ligand concentration. Overall these evidences open new horizons on understanding the influence of conformational binding modes in protein-ligand interactions, which is of paramount importance in rational drugs-design.

- For the first time a combination of molecular docking and DFT results of the pharmacodynamics DDI between two beta-blockers acebutolol and propranolol with relevant catalytic binding-sites of the fibrinogen E-region was proposed. Those results showed that the three most relevant fibrinogen E-region binding sites exhibit from high to low druggability degree in the order: site 1 (D_g = 0.81) > site 2 (D_g = 0.54) > site 3 (D_g = 0.39), as well as complying with the *mutual overlap criterion* for being the best-ranked binding-sites, in contrast with the "worst fibrinogen binding site" with D_g = 0.08. Furthermore, the binding free energy trajectories and total affinity obtained for the best-ranked beta-blocker drug-drug docking complexes reveal a spontaneous thermodynamic stabilization following the site-affinity order: site 1 > site 2 > site 3. The most dominant energy contributions to the total affinity of the complexes (beta-blocker drug-drug/fibrinogen E-region) were provided by the London dispersion attractive forces.

  In addition, the 3D-lig-plot interaction diagrams coupled with DFT results for the best-ranked drug-drug configuration binding-pose (configuration XII) showed that the complex drug-drug stabilization was mostly influenced by the contributions of a high number of hydrophobic atoms (acebutolol > propranolol) in the fibrinogen site 1 (tunnel 2). Besides, the later site was identified as the most preferred biophysical environment for beta-blockers DDI involving all chains (N, O, P, Q, R, S) of the fibrinogen E-region structure and target residues with high depth-values. The obtained evidences may open new perspectives in computational polypharmacology modelling of DDI, which can help avoiding the occurrence of potential undesired effects as perturbations in the coagulation processes induced by concomitant administration of beta-blockers, where the fibrinogen thrombin binding-domain (central E-region) plays a critical role in the context of polypharmacology drug-drug interactions.

- The conformational binding mechanisms of fibrinogen under interaction with recognized β-lactam antibiotics were evaluated. The computational results pointing that despite these conformational differences, both antibiotics exhibit very similar affinity-based on the obtained free binding energy values. Besides, the semi-synthetic incorporation of an additional halogen Cl-atom in the dicloxacillin molecule respect to cloxacillin molecule, and its relative docking-pose orientation in the fibrinogen E-region could significantly reduce the appearance of potential fibrinolytic side-effects affecting the clotting process which are usually associated to β-lactam antibiotic parenterally administered. The performed 3D/2D-lig-plot diagrams revealed that the most relevant antibiotic interactions with the E-region (pocket 1) are mainly based on hydrophobic (C⋯C)-backbone-side-chain non-covalent and acceptor/donor interactions with critical regulatory E-region residues (SER50:Q > SER50:N) with high H-bond atom energy contribution required for the docking complex stabilization for both drugs. Besides, collective low-frequency normal modes and LPRS maps revealed the subtle differences in the conformational dynamics of relevant E-region chains under unbound and bound state with the β-lactam antibiotics. Next, we strongly suggest that the differences in the conformational binding mechanism for dicloxacillin and cloxacillin, predominantly involve the allosteric modulation, long-distance perturbations, and remarkable conformational rigidification of the regulatory chains from E-region (N and Q). Besides, the relevant role that the penicillin molecular structure plays in the conformational binding process with the fibrinogen protein, it has been previously confirmed by our group using experimental validation based on isothermal titration calorimetry (ITC) and quantitative differential scanning calorimetry (DSC).

  Both methods, excellently corroborated the computational predictions from: stoichiometry, target-residues, and binding energies point of view toward "*de novo rational drug-design"* of new β-lactam antibiotic drugs to prevent potential fibrinolytic side-effects, increase the target specificity with optimal benefit/risk rates and contribute to accurate implementation of rigorous criteria for a more personalized antibiotic therapy.

- Structure-based docking and molecular dynamic simulations were combined to study the mitochondrial channel nanotoxicity (ATP transport perturbations in hVDAC1 channel) induced by carbon nanotubes (z-z-SWCNT). The results show that z-z-SWCNT can directly block the ATP-translocation through hVDAC1 channel from the first target-residue (MET1). Besides, the best docking pose of the simulated z-z-SWCNT ligand shows that the interactions with ATP-entry-point are mainly based on van der Waals hydrophobic interactions, revealing very close relative position from the native substrate (ATP molecule).

  Besides, molecular dynamics results fully match with the previous docking attainments, indicating a weakening in the binding affinity of the ATP molecule when the z-z-SWCNT molecule simultaneously interacts in the same biophysical environment (ATP-entry-point). We strongly suggest that the potential mitochondrial channel nanotoxicity (ATP transport dysfunction) of z-z-SWCNT occurs by inducing local perturbations in the inter-residue network topology and transduction properties of the ATP-entry-point. Finally, the obtained results open new opportunities towards modeling the ATP bioenergetic dysfunction induced by carbon nanomaterials interactions, and new strategies of drug discovery-based nanotechnology for precision mitochondrial nanomedicine.

- A combination of experimental and computational approaches to tackle the nanotoxicity of pristine and oxidized single-walled carbon nanotubes (SWCNT-pristine, SWCNT-COOH) based on the mitochondrial F0F1-ATPase inhibition was addressed. Experimental evidences supported that the *in vitro* F0F1-ATPase inhibition responses in submitochondrial particles (SMP) are strongly dependent on the higher level of concentration (from 3 to 5 µg/mL) in both types of single-walled carbon nanotubes (SWCNT-pristine and SWCNT-COOH) evaluated. In addition, both types of carbon nanotubes show an interaction inhibition pattern for the F0F1-ATPase enzyme, similar to the oligomycin A (specific F0F1-ATPase inhibitor). In addition, results on elastic network models (LPRS maps) show that the SWCNT-pristine can promote an abrupt perturbation in several blocks of residues (strong F0-ATPase nanotoxicity inhibition), more pronounced than the analogous SWCNT-COOH (moderate F0-ATPase nanotoxicity inhibition), triggering perturbations on the allosteric responses and abnormal signals' propagation across the inter-residue network of the F0F1-ATPase. In accordance with this, results on the fractal surface

of interactions based on the formed docking complexes (SWCNT-pristine/F0F1-ATPase >> SWCNT-COOH/F0F1-ATPase) suggest that the SWCNT-pristine interactions topologically affect the van der Waals fractal surface and geometric properties of F0-ATPase compared to physiological condition (unbound F0-ATPase). We suggest that the SWCNT-pristine perturbations could negatively impact on catalytic function of F0-ATPase (mitochondrial ATP-hydrolysis), by irreversibly affecting the structural properties of the binding cavities in the F0-subunit. Lastly, the performed Nano-QSBR/QSTR models allow to predict the mitochondrial nanotoxicity (F0-ATPase nanotoxicity inhibition) which was strongly associated to electro-topological and graph-theoretic nanodescriptors from both, SWCNT-pristine and SWCNT-COOH (as Narumi geometric topological index (GNAR) and the electro-topological positive variation (MAXDP)). These results open new opportunities for the best understanding of structural attribute underlying mitochondrial nanotoxicity mechanisms, enhance a rational drug-discovery based on nanotechnology, and boosting precision nanomedicine applications.

- The fullerol nanosystem to be proposed as nanobiomarker early diagnosis of Alzheimer's disease (AD) was computationally modeled and the binding affinity of the nanostructured ligands with the PHF-tau protein was evaluated using a combination of molecular docking and DFT simulation. The in silico results suggest that, all the nano-ligands studied present good affinity with optimal inter-atomic interaction distances, suggesting high specificity and selectivity for tau protein for AD detection purposes. Besides, the evaluated nanobiomarker ligands demonstrated a predominance of hydrophobic and hydrogen bond contacts with the AD tau protein; which is considered as relevant target for molecular recognition purposes for potential biomedical applications. The best docking complex based on the binding affinity by PHF-tau protein was obtained for the F-THK5105 plus fullerol, with a binding energy value of -7.0 kcal/mol and R.M.S.D value of 1.491 Å. In this case, DFT approach allows to obtain the best interface of interaction of the nanobiomarker system (i.e., tau-(Lys340) interacting with atoms from the PHF-18F-THK5105 plus fullerol) shows no significant changes its electronic and structural properties, pointing a physical adsorption regime of interactions. All the arrangements studied strongly suggest that the obtained nanosystems present optimal features of complementarity with AD-tau protein to

be theoretically considered as promising for the sensing and molecular recognition in the context of the early detection of Alzheimer's disease.

- For the first time, new mechanistic approaches-based on molecular docking and DFT-simulations were performed to model the environmental co-exposure of a high concentration of a heterogeneous mixture of oil-derived hydrocarbons (methylnaphthalene, n-butylbenzene, and n-octane) that present a high environmental impact by considering the simultaneous interaction with the detoxifying protein (TodX) from *Pseudomonas putida* and graphene nanostructures.

  The obtained results performing a flexible docking coupled with 3D-lig-plot diagrams allowed mimicking relevant interactions of each oil-derived hydrocarbon (methylnaphthalene, n-butylbenzene, and n-octane) in low-concentrations. For this instance, the Gibbs free energy of binding of the formed complexes (oil-derived hydrocarbon + TodX) exhibit a spontaneous thermodynamic processes ($\Delta G_{bind} < 0$ kcal/mol), with prevalence of non-covalent hydrophobic interactions in the TodX-binding site 1 from with an obtained average FEB-values around ≈ - 25.35 kcal/mol from the three hydrocarbon.

  On the other hand, the graphene nanostructures (pristine-Gr and oxidized-Gr) interact in a different biophysical environment (TodX-binding site 2); describing spontaneous thermodynamic processes following the order FEB (TodX/pristine-Gr complex) = - 26.332 kcal/mol > FEB (TodX/oxidized-Gr complex) = -20.132 kcal/mol according to the strength of the interactions. In addition, the graphene interactions in the TodX-binding site 2 can modulate the responses on the TodX-binding site 1, affecting the communication efficiency in the ToDX residue-network. Further modeling results using heterogeneous multicomponent docking approach (oil-derived hydrocarbons + graphene nanostructures + TodX protein) and DFT-simulation, show a significant decrease of the total affinity of overall the oil-derived hydrocarbons by the TodX (site 1) with an obtained average FEB-values around ≈ - 5.8 kcal/mol compared with the affinity of the complex mimicking low-concentrations FEB-values (oil-derived hydrocarbon + TodX) = - 25.35 kcal/mol from the three hydrocarbons (methylnaphthalene, n-butylbenzene, and n-octane). The general DFT-results suggest the occurrence of physical adsorption mechanism in the reactive surface of the graphene nanostructures

(pristine-Gr > oxidized-Gr), mimicking a *Trojan-horse phenomenon*, which is well-recognized leading the most interactions for organic nanomaterials under co-exposure with complex mixtures.

These theoretical results open new horizons to improve bioremediation strategies to prevent the negative impact of multiple oil-derived hydrocarbons under environmental co-exposure and to ensure a safe and sustainable use of the graphene nanomaterials-based nanotechnologies applications.

The current doctoral thesis focuses on understanding the thermodynamic events of protein-ligand interactions which have been of paramount importance from traditional Medicinal Chemistry to Nanobiotechnology. Particular attention has been made on the application of state-of-the-art methodologies to address thermodynamic studies of the protein-ligand interactions by integrating structure-based molecular docking techniques, classical fractal approaches to solve protein-ligand complementarity problems, perturbation models to study allosteric signal propagation, predictive nano-quantitative structure-toxicity relationship models coupled with powerful experimental validation techniques. The contributions provided by this work could open an unlimited horizon of possibilities for efficient drug-discovery applications. Also, boosting new and suitable strategies to investigate the molecular diversity of ligands and their “bioactivity” properties at very early stages of the preclinical investigation having a direct relevance to the fields of Drug-Discovery, Materials Sciences, Molecular Diagnosis, and Environmental Health Sciences.